

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> A01N 63/00, A61K 39/395, C12N 15/00, A01N 61/00, C07H 21/02	<b>A1</b>	<b>(11) International Publication Number:</b> WO 99/49735 <b>(43) International Publication Date:</b> 7 October 1999 (07.10.99)
<b>(21) International Application Number:</b> PCT/US99/06644 <b>(22) International Filing Date:</b> 26 March 1999 (26.03.99)  <b>(30) Priority Data:</b> 60/079,759 27 March 1998 (27.03.98) US 60/095,153 3 August 1998 (03.08.98) US  <b>(71) Applicant (for all designated States except US):</b> FOX CHASE CANCER CENTER [US/US]; 7701 Burholme Avenue, Philadelphia, PA 19111 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> KRUH, Gary [US/US]; 241 South 6th Street #809, Philadelphia, PA 19106 (US). LEE, Kun [KR/US]; 21 Barrington Drive, Cranbury, NJ 08512 (US). BELINSKY, Martin [US/US]; 625 Parmentier Road, Warminster, PA 18974 (US). BAIN, Lisa [US/US]; 284 Penny Lane, Townville, SC 29689 (US).  <b>(74) Agents:</b> RIGAUT, Kathleen, D. et al.; Dann, Dorfman, Herrell and Skillman, Suite 720, 1601 Market Street, Philadelphia, PA 19103 (US).		<b>(81) Designated States:</b> AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> MPR-RELATED ABC TRANSPORTER ENCODING NUCLEIC ACIDS AND METHODS OF USE THEREOF		
<b>(57) Abstract</b>  Novel human MOAT genes and their encoded proteins are provided herein. The MRP-related ABC transporters encoded by the disclosed nucleic acid sequences play a pivotal role in the efflux of pharmacologically beneficial reagents from tumor cells. MOAT genes and their encoded proteins provide valuable therapeutic targets for the design of anti-cancer agents which inhibit the aberrant growth of malignant cells.		

BEST AVAILABLE COPY

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakistan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**MRP-Related ABC Transporter  
Encoding Nucleic Acids and Methods of Use Thereof**

Pursuant to 35 U.S.C. §202(c) it is acknowledged that the U.S. Government has certain rights in the invention described herein, which was made in part with funds from the National Institutes of Health, Grant Numbers, CA63173 and CA06927.

**FIELD OF THE INVENTION**

The present invention relates to the fields of medicine and molecular biology. More specifically, the invention provides nucleic acid molecules and proteins encoded thereby which are involved in the development of resistance to pharmacological and chemotherapeutic agents in tumor cells.

**BACKGROUND OF THE INVENTION**

Several publications are referenced in this application in parentheses in order to more fully describe the state of the art to which this invention pertains. The disclosure of each of these publications is incorporated by reference herein.

P-glycoprotein, the product of the *MDR1* gene, was the first ABC transporter shown to confer resistance to cytotoxic agents. Pgp functions as an ATP-dependent efflux pump that reduces the intracellular concentration of a variety of chemotherapeutic agents by transporting them across the plasma membrane (1). The multidrug resistance phenotype associated with overexpression of Pgp

is of considerable clinical interest because natural product drugs are second only to alkylating agents in clinical utility, and many effective chemotherapeutic regimens contain more than one natural product agent. More recently, we and others have reported transfection studies indicating that MRP, another ABC family transporter, confers a multidrug resistance phenotype that includes many natural product drugs, but is distinct from the resistance phenotype associated with Pgp (2-6). MRP shares only limited amino acid identity with Pgp, and this is reflected in the different substrate specificities of the two transporters. In contrast to Pgp, MRP can transport a wide range of anionic organic conjugates, including glutathione S-conjugates (7). In addition to Pgp and MRP there may be other transporters that are involved in cytotoxic drug resistance. In the case of natural product drugs, resistant cell lines have been described that display a multidrug resistant phenotype associated with a drug accumulation deficit, but do not overexpress Pgp or MRP (8). ABC transporters have also been linked to cisplatin resistance, and several lines of evidence suggest the possibility that pumps specific for organic anions may be involved: 1) decreased cisplatin accumulation is consistently observed in cisplatin resistant cell lines (9); 2) cisplatin is conjugated to glutathione in the cell, and this anionic conjugate is toxic in an *in vitro* biochemical assay (10); and 3) biochemical studies using membrane vesicle preparations have shown that cisplatin resistant cells lines have enhanced expression of an ATP-dependent transporter of CDDP-glutathione and other glutathione S-conjugates such as the cystinyl leukotriene LTC<sub>4</sub> (11, 12). These data thus suggest that an organic anion transporter may contribute

to cisplatin resistance by exporting CDDP-glutathione. While MRP is an organic anion transporter, the reported drug resistance profile of MRP-transfected cells does not extend to this agent (5, 6), and to date only one cisplatin resistant cell line has been reported to overexpress MRP (13). This suggests that organic anion transporters other than MRP may contribute to cisplatin resistance. Consistent with this possibility, the canalicular multispecific organic anion transporter, cMOAT, an MRP-related transporter that functions as the major organic anion transporter in liver, has been reported to be overexpressed in cisplatin resistant cell lines (14, 15). A more direct link between cMOAT and cytotoxic drug resistance is suggested by a recent report in which transfection of a cMOAT antisense construct into a liver cancer cell line resulted in sensitization to cisplatin, daunorubicin and other cytotoxic agents (16).

Clearly, a need exists for identifying the essential components and mechanisms giving rise to drug resistance and the transport of anticancer agents out of the tumor cell. The elucidation of these mechanisms may be used to advantage for the design of efficacious chemotherapeutic agents.

#### **SUMMARY OF THE INVENTION**

This invention provides novel, biological molecules useful for identification, detection, and/or molecular characterization of components involved in the acquisition of drug resistance in tumor cells. According to one aspect of the invention, an isolated nucleic acid molecule is provided which includes a sequence encoding a protein transporter of a size between about 1300 and 1350 amino acids in length. The encoded protein, referred to herein

as MOAT-B, comprises a multi-domain structure including a tandem repeat of nucleotide binding folds appended C-terminal to a hydrophobic domain that contains several potential membrane spanning helices. Conserved Walker A and B ATP binding sites are present in each of the nucleotide binding folds.

In a preferred embodiment of the invention, an isolated nucleic acid molecule is provided that includes a cDNA encoding a human MOAT-B protein. In a particularly preferred embodiment, the human MOAT-B protein has an amino acid sequence the same as Sequence I.D. No. 2. An exemplary MOAT-B nucleic acid molecule of the invention comprises Sequence I.D. No. 1.

According to another aspect of the invention, a second isolated nucleic acid molecule is provided which includes a sequence encoding a transporter between about 1400 and 1450 amino acids. The encoded protein, referred to herein as MOAT-C contains a multi-domain structure including a tandem repeat of nucleotide binding folds appended C-terminal to a hydrophobic domain that contains several potential membrane spanning helices. Conserved Walker A and B ATP binding sites are present in each of the nucleotide binding folds. While similar in structure to MOAT-B described above, MOAT-C contains distinct sequence differences.

In a preferred embodiment of the invention, an isolated nucleic acid molecule is provided that includes a cDNA encoding a human MOAT-C protein. In a particularly preferred embodiment, the human MOAT-C protein has an amino acid sequence the same as Sequence I.D. No. 4. An exemplary MOAT-C nucleic acid molecule of the invention comprises Sequence I.D. No. 3.

According to yet another aspect of the invention, an

isolated nucleic acid molecule is provided which includes a sequence encoding a protein of a size between about 1500 and 1550 amino acids in length. The encoded protein, referred to herein as MOAT-D, contains a multidomain structure including an N-terminal hydrophobic extension which harbors five transmembrane spanning helices.

In a preferred embodiment of the invention, an isolated nucleic acid molecule is provided that includes a cDNA encoding a MOAT-D protein. In a particularly preferred embodiment, the human MOAT-D protein has an amino acid sequence the same as Sequence I.D. No. 6. An exemplary MOAT-D nucleic acid molecule of the invention comprises Sequence I.D. No. 5.

According to yet another aspect of the invention, an isolated nucleic acid molecule is provided which includes a sequence encoding a protein of a size between about 1480 and 1530 amino acids in length. The encoded protein, referred to herein as MOAT-E, contains a multidomain structure including an N-terminal hydrophobic extension which harbors several transmembrane spanning helices. While similar in structure to MOAT-D described above, MOAT-E contains distinct sequence differences.

In a preferred embodiment of the invention, an isolated nucleic acid molecule is provided that includes a cDNA encoding a MOAT-E protein. In a particularly preferred embodiment, the human MOAT-E protein has an amino acid sequence the same as Sequence I.D. No. 8. An exemplary MOAT-E nucleic acid molecule of the invention comprises Sequence I.D. No. 7.

According to another aspect of the present invention, an isolated nucleic acid molecule is provided, which has a sequence selected from the group consisting of: (1) Sequence I.D. No. 1; (2) a sequence specifically

hybridizing with preselected portions or all of the complementary strand of Sequence I.D. No. 1 comprising nucleic acids encoding amino acids 1-1154 of Sequence ID No. 2; (3) a sequence encoding preselected portions of Sequence I.D. No. 1 within nucleotides 1-3462, (4) Sequence I.D. No. 3; (5) a sequence specifically hybridizing with preselected portions or all of the complementary strand of Sequence I.D. No. 3 comprising nucleic acids encoding amino acids 1-442 of Sequence ID No. 4; (6) a sequence encoding preselected portions of Sequence I.D. No. 3 within nucleotides 1-1326, (7) Sequence I.D. No. 5; (8) a sequence specifically hybridizing with preselected portions or all of the complementary strand of Sequence I.D. No. 5 comprising nucleic acids encoding amino acids 1-1036 of Sequence ID No. 6; (9) a sequence encoding preselected portions of Sequence I.D. No. 5 within nucleotides 1-3108, (1) Sequence I.D. No. 7; (2) a sequence specifically hybridizing with preselected portions or all of the complementary strand of Sequence I.D. No. 7 comprising nucleic acids encoding amino acids 1-998 of Sequence ID No. 8; (3) a sequence encoding preselected portions of Sequence I.D. No. 7 within nucleotides 1-300.

Such partial sequences are useful as probes to identify and isolate homologues of the MOAT genes of the invention. Additionally, isolated nucleic acid sequences encoding natural allelic variants of the nucleic acids of Sequence I.D. Nos., 1, 3, 5 and 7 are also contemplated to be within the scope of the present invention. The term natural allelic variants will be defined hereinbelow.

According to another aspect of the present invention, antibodies immunologically specific for the human MOAT proteins described hereinabove are provided.



In yet another aspect of the invention, host cells comprising at least one of the MOAT encoding nucleic acids are provided. Such host cells include but are not limited to bacterial cells, fungal cells, insect cells, mammalian cells, and plant cells. Host cells overexpressing one or more of the MOAT encoding nucleic acids of the invention provide valuable research tools for assessing transport of chemotherapeutic agents out of cells. MOAT expressing cells also comprise a biological system useful in methods for identifying inhibitors of the MOAT transporters.

Another embodiment of the present invention encompasses methods for screening cells expressing MOAT encoding nucleic acids for chemotherapy resistance. Such methods will provide the clinician with data which correlates expression of a particular MOAT genes with a particular chemotherapy resistant phenotype.

Diagnostic methods are also contemplated in the present invention. Accordingly, suitable oligonucleotide probes are provided which hybridize to the nucleic acids of the invention. Such probes may be used to advantage in screening biopsy samples for the expression of particular MOAT genes. Once a tumor sample has been characterized as to the MOAT gene(s) expressed therein, inhibitors identified in the cell line screening methods described above may be administered to prevent efflux of the beneficial chemotherapeutic agents from cancer cells.

The methods of the invention may be applied to kits. An exemplary kit of the invention comprises MOAT gene specific oligonucleotide probes and/or primers, MOAT encoding DNA molecules for use as a positive control, buffers, and an instruction sheet. A kit for practicing the cell line screening method includes frozen cells

comprising the MOAT genes of the invention, suitable culture media, buffers and an instruction sheet.

In a further aspect of the invention, transgenic knockout mice are disclosed. Mice will be generated in which at least one MOAT gene has been knocked out. Such mice will provide a valuable in biological system for assessing resistance to chemotherapy in an in vivo tumor model.

Various terms relating to the biological molecules of the present invention are used hereinabove and also throughout the specification and claims. The terms "percent similarity" and "percent identity (identical)" are used as set forth in the UW GCG Sequence Analysis program (Devereux et al. NAR 12:387-397 (1984)).

With reference to nucleic acids of the invention, the term "isolated nucleic acid" is sometimes used. This term, when applied to DNA, refers to a DNA molecule that is separated from sequences with which it is immediately contiguous (in the 5' and 3' directions) in the naturally occurring genome of the organism from which it originates. For example, the "isolated nucleic acid" may comprise a DNA or cDNA molecule inserted into a vector, such as a plasmid or virus vector, or integrated into the genomic DNA of a prokaryote or eukaryote.

With respect to RNA molecules of the invention, the term "isolated nucleic acid" primarily refers to an RNA molecule encoded by an isolated DNA molecule as defined above. Alternatively, the term may refer to an RNA molecule that has been sufficiently separated from RNA molecules with which it would be associated in its natural state (i.e., in cells or tissues), such that it exists in a "substantially pure" form (the term "substantially pure" is defined below).

With respect to protein, the term "isolated protein" or "isolated and purified protein" is sometimes used herein. This term refers primarily to a protein produced by expression of an isolated nucleic acid molecule of the invention. Alternatively, this term may refer to a protein which has been sufficiently separated from other proteins with which it would naturally be associated, so as to exist in "substantially pure" form.

The term "substantially pure" refers to a preparation comprising at least 50-60% by weight the compound of interest (e.g., nucleic acid, oligonucleotide, protein, etc.). More preferably, the preparation comprises at least 75% by weight, and most preferably 90-99% by weight, the compound of interest. Purity is measured by methods appropriate for the compound of interest (e.g. chromatographic methods, agarose or polyacrylamide gel electrophoresis, HPLC analysis, and the like). With respect to antibodies of the invention, the term "immunologically specific" refers to antibodies that bind to one or more epitopes of a protein of interest (e.g., MOAT-B, MOAT-C or MOAT-D), but which do not substantially recognize and bind other molecules in a sample containing a mixed population of antigenic biological molecules.

With respect to nucleic acids and oligonucleotides, the term "specifically hybridizing" refers to the association between two single-stranded nucleotide molecules of sufficiently complementary sequence to permit such hybridization under pre-determined conditions generally used in the art (sometimes termed "substantially complementary"). When used in reference to a double stranded nucleic acid, this term is intended to signify that the double stranded nucleic acid has been subjected to denaturing conditions, as is well known to those of

skill in the art. In particular, the term refers to hybridization of an oligonucleotide with a substantially complementary sequence contained within a single-stranded DNA or RNA molecule of the invention, to the substantial exclusion of hybridization of the oligonucleotide with single-stranded nucleic acids of non-complementary sequence.

One common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid molecules of a specified sequence homology (Sambrook et al., 1989):

$$T_m = 81.5^{\circ}\text{C} + 16.6\text{Log} [\text{Na}^+] + 0.41(\% \text{ G+C}) - 0.63 (\% \text{ formamide}) - 600/\text{\#bp in duplex}$$

As an illustration of the above formula, using  $[\text{Na}^+] = [0.368]$  and 50% formamide, with GC content of 42% and an average probe size of 200 bases, the  $T_m$  is  $57^{\circ}\text{C}$ . The  $T_m$  of a DNA duplex decreases by 1 -  $1.5^{\circ}\text{C}$  with every 1% decrease in homology. Thus, targets with greater than about 75% sequence identity would be observed using a hybridization temperature of  $42^{\circ}\text{C}$ . Such sequences would be considered substantially homologous to the nucleic acid sequences of the invention.

The nucleic acids, proteins, antibodies, cell lines, methods, and kits of the present invention may be used to advantage to identify targets for the development of novel agents which inhibit the aberrant transport of cytotoxic agents out of tumor cells. The transgenic mice of the invention may be used as an in vivo model for chemotherapy resistance.

The human MOAT molecules methods and kits described above may also be used as research tools and will facilitate the elucidation of the mechanism by which tumor

cells acquire a drug resistant phenotype.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the predicted structure of MOAT-B and comparison with human MRP. The vertical lines indicate identical amino acids and the vertical dots indicate conserved amino acids. Gaps are indicated by periods. The overbars indicate potential transmembrane spanning segments as predicted by the TMAP program. The first and second nucleotide binding folds (NBF 1 and NBF 2) are indicated by horizontal arrows. The C-terminal 34 amino acids (residues 1291 - 1325) are replaced in the second class of MOAT-B cDNA clones by the following amino acids: ILQKKLSTYWSH. The Alignment was performed using the GAP program (gap weight 3.0, length weight 0.1) in the Genetics Computer Group Package. H. MRP: human MRP.

Figures 2A and 2B depict a comparison of the nucleotide binding folds and hydropathy profile of MOAT-B with those of other eukaryotic ABC transporters. Fig. 1A shows the comparison of the nucleotide binding folds of MOAT-B. Amino acids that are identical to those of MOAT-B are shaded, and gaps are indicated by periods. Walker A and B motifs, and the ABC transporter family signature sequence C, are underlined. Amino acid positions are indicated to the right. Amino acid sequences were aligned using the PILEUP program (gap weight 3.0, length weight 0.1) in the Genetics Computer Group Package. Fig. 1B shows a comparison of the MOAT-B hydropathy profile. To facilitate comparison, the proteins are aligned so that the N-terminal nucleotide binding folds (NBF) are roughly in register. NBF's are indicated by bars. Values above

and below the horizontal lines indicate hydrophobic and hydrophilic regions, respectively. Hydrophobicity plots were generated using the Kyte-Doolittle algorithm with a window of 7 residues. The transporters shown are: human multidrug-associated protein, H. MRP (P33529); human multispecific organic anion transporter, H. MOAT (U63970); *Saccharomyces cerevisiae* yeast cadmium factor 1, S. YCF1 (P39109); rat sulfonylurea receptor, R. SUR (Q09427); human cystic fibrosis transmembrane conductance regulator, H. CFTR (M28668); *Leishmania* P-glycoprotein, L. PgpA (P21441) and human *mdr1* gene product, H. MDR1 (P08183). Accession numbers are shown in parentheses.

Figure 3 is a Northern blot showing the tissue distribution of MOAT-B transcript. Membranes containing poly (A)+ RNA prepared from human tissues were hybridized with a radiolabeled MOAT-B or GAPDH probe. Top panels show MOAT-B transcript and bottom panels show the control GAPDH transcript. Arrows indicate the position of MOAT-B transcript. Prolonged exposure of the film revealed a low level signal in liver.

Figure 4 shows the chromosomal localization of the gene encoding MOAT-B. Human metaphase spreads were hybridized with a biotin-labeled MOAT-B cDNA probe and detected by FITC-conjugated avidin. Hybridization signals at chromosome 13q32 in two metaphase spreads are indicated by arrows. The inset shows paired hybridization signals at band q32 of chromosome 13 from three other metaphase spreads.

Figures 5A and 5B show the predicted structures of MOAT-C and MOAT-D. Fig. 5A presents the structure of

MOAT-C. Fig. 5B shows the structure of MOAT-D. Numbered overbars indicate potential transmembrane spanning helices. Horizontal arrows indicate the positions of the amino terminal (NBF1) and C-terminal (NBF2) nucleotide binding folds. Walker A and B motifs, and the ABC transporter family signature sequence C are underlined. Bullets indicate the positions of potential N-linked glycosylation sites that are conserved with previously reported N-glycosylation sites in MRP. The indicated MOAT-C transmembrane spanning helices were predicted using the TMAP program and an input alignment of MOAT-B and MOAT-C. The indicated MOAT-D transmembrane helices are based upon inspection of an alignment with MRP.

Figures 6A and 6B show a comparison of the nucleotide binding folds and hydropathy profiles of MOAT-C and MOAT-D with those of other related ABC transporters. Fig. 6A depicts the comparison of the nucleotide binding folds. The alignment was produced using the PILEUP command (gap weight 3.0, length weight 0.1) in the Genetics Computer Group Package Version 9.1. Amino acid positions conserved in at least 4 of the 8 proteins are shaded. Periods indicate gaps in the alignment. Walker A and B, and the ABC transporter family signature sequence C are indicated by underbars. Fig. 6A shows the comparison of hydropathy profiles. To facilitate comparisons, gaps were introduced at the N-termini of some proteins in order to bring the first nucleotide binding folds into register. Nucleotide binding folds are indicated by bars. Values above and below the horizontal lines indicate hydrophobic and hydrophilic regions, respectively. Hydrophobicity plots were generated using the Kyte-Doolittle algorithm with a window of 7 residues. Accession numbers are as follows:

MRP, P33529; cMOAT, U63970, SUR, Q09428; CFTR, P-13569; MDR1, P08183.

Figure 7 is a Northern blot showing the tissue distribution of MOAT-C and MOAT-D transcripts. Blots containing poly A+ RNA prepared from various human tissues were hybridized with MOAT-C, MOAT-D and actin probes. Arrows indicate the position of the MOAT-C (top panel) and MOAT-D (middle panel) transcripts. The bottom panel shows the control actin transcript.

Figures 8A and 8B show the chromosomal localization of the *MOAT-C* and *MOAT-D* genes. Human metaphase spreads were hybridized with a biotin-labeled *MOAT-C* and *MOAT-D* cDNA probes and detected by FITC-conjugated avidin. Fig. 8A shows the localization of *MOAT-C*. Hybridization signals at chromosome 3q27 in two metaphase spreads are indicated by arrows (top). The inset shows paired hybridization signals at band q27 of chromosome 3 from three other metaphase spreads. Fig. 8B shows the localization of *MOAT-D*. Hybridization signals at chromosome 17q21-22 in two metaphase spreads are indicated by arrows (top). The inset shows paired hybridization signals at band q21-22 of chromosome 17 from three other metaphase spreads.

Figure 9 shows predicted amino acid sequence of MOAT-E. Also shown are the location of the potential transmembrane helices (overbars), the potential N-glycosylation site (black dot) and the two nucleotide binding folds (NBF1 and NBF2). Walker A and B motifs, as well as the signature C motif of ABC transporters, are also indicated.



Figure 10 shows a comparison of the hydropathy profile of MOAT-E with other members of the MRP-cMOAT subfamily. The profile reveals that MOAT-E has a hydrophobic N-terminal segment which is absent in MOAT-B and MOAT-C.

Figure 11 is a RNA blot which reveals that MOAT-E is expressed only in the liver and the kidney, suggesting that MOAT-E may participate in the excretion of substances into urine and bile. The lower panel shows hybridization of an actin probe to assess RNA loading.

Figures 12A-12J show the cDNA (SEQ ID NO: 1) and amino acid sequences (SEQ ID NO: 2) encoded by MOATB.

Figures 13A-13K show the cDNA (SEQ ID NO: 3) and amino acid sequences (SEQ ID NO: 4) encoded by MOATC.

Figures 14A-14K show the cDNA (SEQ ID NO: 5) and amino acid sequences (SEQ ID NO: 6) encoded by MOATD.

Figures 15A-15K show the cDNA (SEQ ID NO: 7) and amino acid sequences (SEQ ID NO: 8) encoded by MOATE.

#### **DETAILED DESCRIPTION OF THE INVENTION**

MRP and cMOAT are closely related mammalian ABC transporters that export organic anions from cells. Transfection studies have established that MRP confers resistance to natural product cytotoxic agents, and recent evidence suggests the possibility that cMOAT may contribute to cytotoxic drug resistance as well. Based upon the potential importance of these transporters in

clinical drug resistance, and their important physiological roles in the export of the amphiphilic products of phase I and phase II metabolism, we sought to identify other MRP-related transporters. Using a degenerate PCR approach, a cDNA molecule was isolated which encodes a novel ABC transporter designated herein as MOAT-B. The MOAT-B gene was mapped using fluorescence *in situ* hybridization to chromosome band 13q32. Comparison of the MOAT-B predicted protein with other transporters revealed that it is most closely related to MRP, cMOAT, and the yeast organic anion transporter YCF1. While MOAT-B is closely related to these transporters, it is distinguished by the absence of approximately 200 amino acid N-terminal hydrophobic extension that is present in MRP and cMOAT, and which is predicted to encode several transmembrane spanning segments. In addition, the MOAT-B tissue distribution is distinct from MRP and cMOAT. In contrast to MRP, which is widely expressed in most tissues, including liver, and cMOAT, whose expression is largely restricted to liver, the MOAT-B transcript is widely expressed, with particularly high levels in prostate, but is barely detectable in liver. These data indicate that MOAT-B is a ubiquitously expressed transporter that is closely related to MRP and cMOAT, and indicate that it is an organic anion pump relevant to cellular detoxification.

Three additional MRP/cMOAT-related transporters, MOAT-C, MOAT-D and MOAT-E are also disclosed herein. MOAT-C encodes a 1437 amino acid protein that is most closely related to MRP, cMOAT and MOAT-B, among eukaryotic transporters (33% - 37% identity). However, based upon amino acid identity, MOAT-C is considerably less related to MRP and cMOAT than the latter transporters are to each

other (48% identity). In addition, the MOAT-C topology is distinct from that of MRP and cMOAT in that it, like MOAT-B, lacks an N-terminal transmembrane spanning domain. MOAT-D encodes a 1530 amino acid transporter that is highly related to MRP (57% identity) and cMOAT (47% identity). MOAT-E encodes 1503 amino acid transporter that is highly related to MOAT-D, MRP and cMOAT (39-45% identity). The topology of MOAT-D and MOAT-E are quite similar to MRP and cMOAT, in that they have an N-terminal hydrophobic extension that is predicted to harbor five transmembrane spanning helices. MOAT-C and MOAT-D were mapped to chromosome bands 3q27 and 17q21-22, respectively, by fluorescence *in situ* hybridization.

The expression patterns of MOAT-C, MOAT-D and MOAT-E are distinct from those of MRP, cMOAT and MOAT-B. MOAT-C transcript is widely expressed, with highest levels in skeletal muscle, kidney and testis, but is expressed at barely detectable levels in liver and lung. MOAT-D transcript has a more restricted expression pattern, with high levels in colon, pancreas, liver and kidney. Data presented herein reveal that MOAT-E expression is restricted to liver and kidney.

Based upon degree of amino acid identity, and protein topology, the MRP-related transporters fall into two groups, with the first group consisting of MRP, cMOAT, MOAT-D and MOAT-E, and the second group consisting of MOAT-B and MOAT-C. The isolation of MOAT-C, MOAT-D and MOAT-E thus helps to define the MRP/cMOAT subfamily. The high degree of amino acid identity and topological similarity of MOAT-D and MOAT-E to MRP and cMOAT suggest that they function as organic anion transporters, and play a role in cytotoxic drug resistance. In contrast, the lower degree of amino acid identify and distinct topology

of MOAT-B and MOAT-C suggest the possibility that their substrate specificities and functions may be distinct from that of MRP, cMOAT, MOAT-D and MOAT-E.

The compositions, methods, kits and transgenic mice of the invention disclosed herein will facilitate the identification of drugs that cripple the ability of MOAT genes and proteins encoded thereby to effect the efflux of clinically beneficial pharmacological agents in malignant cells.

**I. Preparation of MOAT-Encoding Nucleic Acid Molecules, MOAT Proteins, and Antibodies Thereto**

**A. Nucleic Acid Molecules**

Nucleic acid molecules encoding the MOAT proteins of the invention may be prepared by two general methods: (1) synthesis from appropriate nucleotide triphosphates, or (2) isolation from biological sources. Both methods utilize protocols well known in the art. The availability of nucleotide sequence information, such as cDNAs having Sequence I.D. Nos. 1, 3, 5, or 7 enables preparation of an isolated nucleic acid molecule of the invention by oligonucleotide synthesis. Synthetic oligonucleotides may be prepared by the phosphoramidite method employed in the Applied Biosystems 38A DNA Synthesizer or similar devices. The resultant construct may be purified according to methods known in the art, such as high performance liquid chromatography (HPLC). Long, double-stranded polynucleotides, such as a DNA molecule of the present invention, must be synthesized in stages, due to the size limitations inherent in current oligonucleotide synthetic methods. Thus, for example, a 5 kb double-stranded molecule may be synthesized as several smaller segments of appropriate complementarity. Complementary segments thus

produced may be annealed such that each segment possesses appropriate cohesive termini for attachment of an adjacent segment. Adjacent segments may be ligated by annealing cohesive termini in the presence of DNA ligase to construct an entire 5 kb double-stranded molecule. A synthetic DNA molecule so constructed may then be cloned and amplified in an appropriate vector.

Nucleic acid sequences encoding the MOAT proteins of the invention may be isolated from appropriate biological sources using methods known in the art. In a preferred embodiment, a cDNA clone is isolated from a cDNA expression library of human origin. In an alternative embodiment, utilizing the sequence information provided by the cDNA sequence, human genomic clones encoding MOAT proteins may be isolated. Alternatively, cDNA or genomic clones having homology with MOAT-B, MOAT-C, MOAT-D or MOAT-E may be isolated from other species using oligonucleotide probes corresponding to predetermined sequences within the MOAT encoding nucleic acids.

In accordance with the present invention, nucleic acids having the appropriate level of sequence homology with the protein coding region of Sequence I.D. Nos. 1, 3, 5, and 7 may be identified by using hybridization and washing conditions of appropriate stringency. For example, hybridizations may be performed, according to the method of Sambrook et al., (supra) using a hybridization solution comprising: 5X SSC, 5X Denhardt's reagent, 1.0% SDS, 100 µg/ml denatured, fragmented salmon sperm DNA, 0.05% sodium pyrophosphate and up to 50% formamide. Hybridization is carried out at 37-42°C for at least six hours. Following hybridization, filters are washed as follows: (1) 5 minutes at room temperature in 2X SSC and 1% SDS; (2) 15 minutes at room temperature in 2X SSC and

0.1% SDS; (3) 30 minutes-1 hour at 37°C in 1X SSC and 1% SDS; (4) 2 hours at 42-65° in 1X SSC and 1% SDS, changing the solution every 30 minutes.

Nucleic acids of the present invention may be maintained as DNA in any convenient cloning vector. In a preferred embodiment, clones are maintained in a plasmid cloning/expression vector, such as pBluescript (Stratagene, La Jolla, CA), which is propagated in a suitable *E. coli* host cell.

MOAT-encoding nucleic acid molecules of the invention include cDNA, genomic DNA, RNA, and fragments thereof which may be single- or double-stranded. Thus, this invention provides oligonucleotides (sense or antisense strands of DNA or RNA) having sequences capable of hybridizing with at least one sequence of a nucleic acid molecule of the present invention, such as selected segments of the cDNA having Sequence I.D. No. 1. Such oligonucleotides are useful as probes for detecting or isolating MOAT genes. Antisense nucleic acid molecules may be targeted to translation initiation sites and/or splice sites to inhibit the translation of the MOAT-encoding nucleic acids of the invention. Such antisense molecules are typically between 15 and 30 nucleotides in length and often span the translational start site of MOAT encoding mRNA molecules.

It will be appreciated by persons skilled in the art that variants of these sequences exist in the human population, and must be taken into account when designing and/or utilizing oligos of the invention. Accordingly, it is within the scope of the present invention to encompass such variants, with respect to the MOAT sequences disclosed herein or the oligos targeted to specific locations on the respective genes or RNA transcripts.

With respect to the inclusion of such variants, the term "natural allelic variants" is used herein to refer to various specific nucleotide sequences and variants thereof that would occur in a human population. The usage of different wobble codons and genetic polymorphisms which give rise to conservative or neutral amino acid substitutions in the encoded protein are examples of such variants. Additionally, the term "substantially complementary" refers to oligo sequences that may not be perfectly matched to a target sequence, but the mismatches do not materially affect the ability of the oligo to hybridize with its target sequence under the conditions described.

#### **B. Proteins**

Full-length MOAT-B, MOAT-C, MOAT-D and MOAT-E proteins of the present invention may be prepared in a variety of ways, according to known methods. The proteins may be purified from appropriate sources, e.g., transformed bacterial or animal cultured cells or tissues, by immunoaffinity purification. However, this is not a preferred method due to the low amount of protein likely to be present in a given cell type at any time. The availability of nucleic acid molecules encoding MOAT proteins enables production of the proteins using *in vitro* expression methods known in the art. For example, a cDNA or gene may be cloned into an appropriate *in vitro* transcription vector, such as pSP64 or pSP65 for *in vitro* transcription, followed by cell-free translation in a suitable cell-free translation system, such as wheat germ or rabbit reticulocytes. *In vitro* transcription and translation systems are commercially available, e.g., from Promega Biotech, Madison, Wisconsin or Gibco-BRL,

Gaithersburg, Maryland.

Alternatively, according to a preferred embodiment, larger quantities of MOAT proteins may be produced by expression in a suitable prokaryotic or eukaryotic system. For example, part or all of a DNA molecule, such as a cDNA having Sequence I.D. No. 1, 3, 5 or 7 may be inserted into a plasmid vector adapted for expression in a bacterial cell, such as *E. coli*. Such vectors comprise the regulatory elements necessary for expression of the DNA in the host cell positioned in such a manner as to permit expression of the DNA in the host cell. Such regulatory elements required for expression include promoter sequences, transcription initiation sequences and, optionally, enhancer sequences.

The human MOAT proteins produced by gene expression in a recombinant procaryotic or eukaryotic system may be purified according to methods known in the art. In a preferred embodiment, a commercially available expression/secretion system can be used, whereby the recombinant protein is expressed and thereafter secreted from the host cell, to be easily purified from the surrounding medium. If expression/secretion vectors are not used, an alternative approach involves purifying the recombinant protein by affinity separation, such as by immunological interaction with antibodies that bind specifically to the recombinant protein or nickel columns for isolation of recombinant proteins tagged with 6-8 histidine residues at their N-terminus or C-terminus. Alternative tags may comprise the FLAG epitope or the hemagglutinin epitope. Such methods are commonly used by skilled practitioners.

The human MOAT proteins of the invention, prepared by the aforementioned methods, may be analyzed according to



standard procedures. For example, such proteins may be subjected to amino acid sequence analysis, according to known methods.

The present invention also provides antibodies capable of immunospecifically binding to proteins of the invention. Polyclonal antibodies directed toward human MOAT proteins may be prepared according to standard methods. In a preferred embodiment, monoclonal antibodies are prepared, which react immunospecifically with the various epitopes of the MOAT proteins described herein. Monoclonal antibodies may be prepared according to general methods of Köhler and Milstein, following standard protocols. Polyclonal or monoclonal antibodies that immunospecifically interact with MOAT proteins can be utilized for identifying and purifying such proteins. For example, antibodies may be utilized for affinity separation of proteins with which they immunospecifically interact. Antibodies may also be used to immunoprecipitate proteins from a sample containing a mixture of proteins and other biological molecules. Other uses of anti-MOAT antibodies are described below.

## **II. Uses of MOAT-Encoding Nucleic Acids, MOAT Proteins and Antibodies Thereto**

Cellular transporter molecules have received a great deal of attention as potential targets of chemotherapeutic agents designed to effectively block the export of pharmacological reagents from tumor cells. The MOAT proteins of the invention play a pivotal role in the transport of molecules across the cell membrane.

Additionally, MOAT nucleic acids, proteins and antibodies thereto, according to this invention, may be used as research tools to identify other proteins that are

intimately involved in the transport of molecules into and out of cells. Biochemical elucidation of molecular mechanisms which govern such transport will facilitate the development of novel anti-transport agents that may sensitize tumor cells to conventional chemotherapeutic agents.

#### A. MOAT-Encoding Nucleic Acids

MOAT-encoding nucleic acids may be used for a variety of purposes in accordance with the present invention. MOAT-encoding DNA, RNA, or fragments thereof may be used as probes to detect the presence of and/or expression of genes encoding MOAT proteins. Methods in which MOAT-encoding nucleic acids may be utilized as probes for such assays include, but are not limited to: (1) *in situ* hybridization; (2) Southern hybridization (3) northern hybridization; and (4) assorted amplification reactions such as polymerase chain reactions (PCR).

The MOAT-encoding nucleic acids of the invention may also be utilized as probes to identify related genes from other animal species. As is well known in the art, hybridization stringencies may be adjusted to allow hybridization of nucleic acid probes with complementary sequences of varying degrees of homology. Thus, MOAT-encoding nucleic acids may be used to advantage to identify and characterize other genes of varying degrees of relation to the MOAT genes of the invention. Such information enables further characterization of transporter molecules which give rise to the chemoresistant phenotype of certain tumors. Additionally, they may be used to identify genes encoding proteins that interact with MOAT proteins (e.g., by the "interaction trap" technique), which should further accelerate

identification of the components involved in the acquisition of drug resistance. The MOAT encoding nucleic acids may also be used to generate primer sets suitable for PCR amplification of target MOAT DNA. Criteria for selecting suitable primers are well known to those of ordinary skill in the art.

Nucleic acid molecules, or fragments thereof, encoding MOAT genes may also be utilized to control the production of MOAT proteins, thereby regulating the amount of protein available to participate in cytotoxic drug efflux. As mentioned above, antisense oligonucleotides corresponding to essential processing sites in MOAT-encoding mRNA molecules may be utilized to inhibit MOAT protein production in targeted cells. Alterations in the physiological amount of MOAT proteins may dramatically affect the ability of these proteins to transport pharmacological reagents out of the cell.

Host cells comprising at least one MOAT encoding DNA molecule are encompassed in the present invention. Host cells contemplated for use in the present invention include but are not limited to bacterial cells, fungal cells, insect cells, mammalian cells, and plant cells. The MOAT encoding DNA molecules may be introduced singly into such host cells or in combination to assess the phenotype of cells conferred by such expression. Methods for introducing DNA molecules are also well known to those of ordinary skill in the art. Such methods are set forth in Ausubel et al. eds., Current Protocols in Molecular Biology, John Wiley & Sons, NY, NY 1995, the disclosure of which is incorporated by reference herein.

The availability of MOAT encoding nucleic acids enables the production of strains of laboratory mice carrying part or all of the MOAT genes or mutated

sequences thereof. Such mice may provide an in vivo model for development of novel chemotherapeutic agents.

Alternatively, the MOAT nucleic acid sequence information provided herein enables the production of knockout mice in which the endogenous genes encoding MOAT-B, MOAT-C, MOAT-D or MOAT-E have been specifically inactivated. Methods of introducing transgenes in laboratory mice are known to those of skill in the art. Three common methods include: 1. integration of retroviral vectors encoding the foreign gene of interest into an early embryo; 2. injection of DNA into the pronucleus of a newly fertilized egg; and 3. the incorporation of genetically manipulated embryonic stem cells into an early embryo.

The alterations to the MOAT gene envisioned herein include modifications, deletions, and substitutions. Modifications and deletions render the naturally occurring gene nonfunctional, producing a "knock out" animal. Substitutions of the naturally occurring gene for a gene from a second species results in an animal which produces an MOAT gene from the second species. Substitution of the naturally occurring gene for a gene having a mutation results in an animal with a mutated MOAT protein. A transgenic mouse carrying the human MOAT gene is generated by direct replacement of the mouse MOAT gene with the human gene. These transgenic animals are valuable for use in vivo assays for elucidation of other medical disorders associated with cellular activities modulated by MOAT genes. A transgenic animal carrying a "knock out" of a MOAT encoding nucleic acid is useful for the establishment of a nonhuman model for chemotherapy resistance involving MOAT regulation.

As a means to define the role that MOAT plays in mammalian systems, mice can be generated that cannot make

MOAT proteins because of a targeted mutational disruption of a MOAT gene.

The term "animal" is used herein to include all vertebrate animals, except humans. It also includes an individual animal in all stages of development, including embryonic and fetal stages. A "transgenic animal" is any animal containing one or more cells bearing genetic information altered or received, directly or indirectly, by deliberate genetic manipulation at the subcellular level, such as by targeted recombination or microinjection or infection with recombinant virus. The term "transgenic animal" is not meant to encompass classical cross-breeding or in vitro fertilization, but rather is meant to encompass animals in which one or more cells are altered by or receive a recombinant DNA molecule. This molecule may be specifically targeted to defined genetic locus, be randomly integrated within a chromosome, or it may be extrachromosomally replicating DNA. The term "germ cell line transgenic animal" refers to a transgenic animal in which the genetic alteration or genetic information was introduced into a germ line cell, thereby conferring the ability to transfer the genetic information to offspring. If such offspring in fact, possess some or all of that alteration or genetic information, then they, too, are transgenic animals.

The alteration or genetic information may be foreign to the species of animal to which the recipient belongs, or foreign only to the particular individual recipient, or may be genetic information already possessed by the recipient. In the last case, the altered or introduced gene may be expressed differently than the native gene.

The altered MOAT gene generally should not fully encode the same MOAT protein native to the host animal and

its expression product should be altered to a minor or great degree, or absent altogether. However, it is conceivable that a more modestly modified MOAT gene will fall within the compass of the present invention if it is a specific alteration.

The DNA used for altering a target gene may be obtained by a wide variety of techniques that include, but are not limited to, isolation from genomic sources, preparation of cDNAs from isolated mRNA templates, direct synthesis, or a combination thereof.

A preferred type of target cell for transgene introduction is the embryonal stem cell (ES). ES cells may be obtained from pre-implantation embryos cultured in vitro. Transgenes can be efficiently introduced into the ES cells by standard techniques such as DNA transfection or by retrovirus-mediated transduction. The resultant transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The introduced ES cells thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal.

One approach to the problem of determining the contributions of individual genes and their expression products is to use isolated MOAT genes to selectively inactivate the wild-type gene in totipotent ES cells (such as those described above) and then generate transgenic mice. The use of gene-targeted ES cells in the generation of gene-targeted transgenic mice is known in the art.

Techniques are available to inactivate or alter any genetic region to a mutation desired by using targeted homologous recombination to insert specific changes into chromosomal alleles. However, in comparison with homologous extrachromosomal recombination, which occurs at a frequency approaching 100%, homologous plasmid-

chromosome recombination was originally reported to only be detected at frequencies between  $10^{-6}$  and  $10^{-3}$ .

Nonhomologous plasmid-chromosome interactions are more frequent occurring at levels  $10^5$ -fold to  $10^2$ -fold greater than comparable homologous insertion.

To overcome this low proportion of targeted recombination in murine ES cells, various strategies have been developed to detect or select rare homologous recombinants. One approach for detecting homologous alteration events uses the polymerase chain reaction (PCR) to screen pools of transformant cells for homologous insertion, followed by screening of individual clones. Alternatively, a positive genetic selection approach has been developed in which a marker gene is constructed which will only be active if homologous insertion occurs, allowing these recombinants to be selected directly. One of the most powerful approaches developed for selecting homologous recombinants is the positive-negative selection (PNS) method developed for genes for which no direct selection of the alteration exists. The PNS method is more efficient for targeting genes which are not expressed at high levels because the marker gene has its own promoter. Non-homologous recombinants are selected against by using the Herpes Simplex virus thymidine kinase (HSV-TK) gene and selecting against its nonhomologous insertion with effective herpes drugs such as gancyclovir (GANC) or (1-(2-deoxy-2-fluoro-B-D arabinofluranosyl)-5-iodouracil, (FIAU). By this counter selection, the number of homologous recombinants in the surviving transformants can be increased.

As used herein, a "targeted gene" or "knock-out" is a DNA sequence introduced into the germline or a non-human animal by way of human intervention, including but not

limited to, the methods described herein. The targeted genes of the invention include DNA sequences which are designed to specifically alter cognate endogenous alleles.

Methods of use for the transgenic mice of the invention are also provided herein. Knockout mice of the invention can be injected with tumor cells or treated with carcinogens to generate carcinomas. Such mice provide a biological system for assessing chemotherapy resistance as modulated by a MOAT gene of the invention. Accordingly, therapeutic agents which inhibit the action of these transporters and thereby prevent efflux of beneficial chemotherapeutic agents from tumor cells may be screened in studies using MOAT knock out mice.

As described above, MOAT-encoding nucleic acids are also used to advantage to produce large quantities of substantially pure MOAT proteins, or selected portions thereof.

#### **B. MOAT Proteins and Antibodies**

Purified full length MOAT proteins, or fragments thereof, may be used to produce polyclonal or monoclonal antibodies which also may serve as sensitive detection reagents for the presence and accumulation of MOAT proteins (or complexes containing MOAT proteins) in mammalian cells. Recombinant techniques enable expression of fusion proteins containing part or all of MOAT proteins. The full length proteins or fragments of the proteins may be used to advantage to generate an array of monoclonal antibodies specific for various epitopes of MOAT proteins, thereby providing even greater sensitivity for detection of MOAT proteins in cells.

Polyclonal or monoclonal antibodies immunologically specific for MOAT proteins may be used in



a variety of assays designed to detect and quantitate the proteins. Such assays include, but are not limited to: (1) flow cytometric analysis; (2) immunochemical localization of MOAT proteins in tumor cells; and (3) immunoblot analysis (e.g., dot blot, Western blot) of extracts from various cells. Additionally, as described above, anti-MOAT antibodies can be used for purification of MOAT proteins and any associated subunits (e.g., affinity column purification, immunoprecipitation).

From the foregoing discussion, it can be seen that MOAT-encoding nucleic acids, MOAT expressing vectors, MOAT proteins and anti-MOAT antibodies of the invention can be used to detect MOAT gene expression and alter MOAT protein accumulation for purposes of assessing the genetic and protein interactions involved in the development of drug resistance in tumor cells.

### C. Methods and Kits Employing the

#### Compositions of the Present Invention

From the foregoing discussion, it can be seen that MOAT-encoding nucleic acids, MOAT-expressing vectors, MOAT proteins and anti-MOAT antibodies of the invention can be used to detect MOAT gene expression and alter MOAT protein accumulation for purposes of assessing the genetic and protein interactions giving rise to chemotherapy resistance in tumor cells.

Exemplary approaches for detecting MOAT nucleic acid or polypeptides/proteins include:

- a) comparing the sequence of nucleic acid in the sample with the MOAT nucleic acid sequence to determine whether the sample from the patient contains mutations; or
- b) determining the presence, in a sample from a patient, of the polypeptide encoded by the MOAT gene and,

if present, determining whether the polypeptide is full length, and/or is mutated, and/or is expressed at the normal level; or

c) using DNA restriction mapping to compare the restriction pattern produced when a restriction enzyme cuts a sample of nucleic acid from the patient with the restriction pattern obtained from normal MOAT gene or from known mutations thereof; or,

d) using a specific binding member capable of binding to a MOAT nucleic acid sequence (either normal sequence or known mutated sequence), the specific binding member comprising nucleic acid hybridizable with the MOAT sequence, or substances comprising an antibody domain with specificity for a native or mutated MOAT nucleic acid sequence or the polypeptide encoded by it, the specific binding member being labelled so that binding of the specific binding member to its binding partner is detectable; or,

e) using PCR involving one or more primers based on normal or mutated MOAT gene sequence to screen for normal or mutant MOAT gene in a sample from a patient.

A "specific binding pair" comprises a specific binding member (sbm) and a binding partner (bp) which have a particular specificity for each other and which in normal conditions bind to each other in preference to other molecules. Examples of specific binding pairs are antigens and antibodies, ligands and receptors and complementary nucleotide sequences. The skilled person is aware of many other examples and they do not need to be listed here. Further, the term "specific binding pair" is also applicable where either or both of the specific binding member and the binding partner comprise a part of a large molecule. In embodiments in which the specific

binding pair are nucleic acid sequences, they will be of a length to hybridize to each other under conditions of the assay, preferably greater than 10 nucleotides long, more preferably greater than 15 or 20 nucleotides long.

In most embodiments for screening for alleles giving rise to chemotherapy resistance, the MOAT nucleic acid in biological sample will initially be amplified, e.g. using PCR, to increase the amount of the analyte as compared to other sequences present in the sample. This allows the target sequences to be detected with a high degree of sensitivity if they are present in the sample. This initial step may be avoided by using highly sensitive array techniques that are becoming increasingly important in the art.

The identification of the MOAT gene and its association with a particular chemotherapy resistance paves the way for aspects of the present invention to provide the use of materials and methods, such as are disclosed and discussed above, for establishing the presence or absence in a test sample of a variant form of the gene, in particular an allele or variant specifically associated with chemotherapy resistance. This may be done to assess the propensity of the tumor to exhibit chemotherapy resistance.

In still further embodiments, the present invention concerns immunodetection methods for binding, purifying, removing, quantifying or otherwise generally detecting biological components. The encoded proteins or peptides of the present invention may be employed to detect antibodies having reactivity therewith, or, alternatively, antibodies prepared in accordance with the present invention, may be employed to detect the encoded proteins or peptides. The steps of various useful immunodetection methods have been

described in the scientific literature, such as, e.g., Nakamura et al. (1987).

In general, the immunobinding methods include obtaining a sample suspected of containing a protein, peptide or antibody, and contacting the sample with an antibody or protein or peptide in accordance with the present invention, as the case may be, under conditions effective to allow the formation of immunocomplexes.

The immunobinding methods include methods for detecting or quantifying the amount of a reactive component in a sample, which methods require the detection or quantitation of any immune complexes formed during the binding process. Here, one would obtain a sample suspected of containing a MOAT gene encoded protein, peptide or a corresponding antibody, and contact the sample with an antibody or encoded protein or peptide, as the case may be, and then detect or quantify the amount of immune complexes formed under the specific conditions.

In terms of antigen detection, the biological sample analyzed may be any sample that is suspected of containing the MOAT antigen, such as a tumor tissue section or specimen, a homogenized tissue extract, an isolated cell, a cell membrane preparation, separated or purified forms of any of the above protein-containing compositions.

Contacting the chosen biological sample with the protein, peptide or antibody under conditions effective and for a period of time sufficient to allow the formation of immune complexes (primary immune complexes) is generally a matter of simply adding the composition to the sample and incubating the mixture for a period of time long enough for the antibodies to form immune complexes with, i.e., to bind to, any antigens present. After this time, the sample-antibody composition, such as a tissue

section, ELISA plate, dot blot or Western blot, will generally be washed to remove any non-specifically bound antibody species, allowing only those antibodies specifically bound within the primary immune complexes to be detected.

In general, the detection of immunocomplex formation is well known in the art and may be achieved through the application of numerous approaches. These methods are generally based upon the detection of a label or marker, such as any radioactive, fluorescent, biological or enzymatic tags or labels of standard use in the art. U.S. Patents concerning the use of such labels include U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241, each incorporated herein by reference. Of course, one may find additional advantages through the use of a secondary binding ligand such as a second antibody or a biotin/avidin ligand binding arrangement, as is known in the art.

In one broad aspect, the present invention encompasses kits for use in detecting expression of MOAT encoding nucleic acids in biological samples, including biopsy samples. Such a kit may comprise one or more pairs of primers for amplifying nucleic acids corresponding to the MOAT gene. The kit may further comprise samples of total mRNA derived from tissues expressing at least one or a subset of the MOAT genes of the invention, to be used as controls. The kit may also comprise buffers, nucleotide bases, and other compositions to be used in hybridization and/or amplification reactions. Each solution or composition may be contained in a vial or bottle and all vials held in close confinement in a box for commercial sale. In a further embodiment, the invention encompasses a kit for use in detecting MOAT proteins in chemotherapy

resistant cancer cells comprising antibodies specific for MOAT proteins encoded by the MOAT nucleic acids of the present invention.

Another aspect of the present invention comprises screening methods employing host cells expressing one or more MOAT genes of the invention. An advantage of having discovered the complete coding sequenced of MOAT B-E is that cell lines that overexpress MOATB C D or E can be generated using standard transfection protocols. Cells that overexpress the complete cDNA will also harbor the complete proteins, a feature that is essential for biological activity of proteins. The overexpressing cell lines will be useful in several ways: 1)The drug sensitivity of overexpressing cell lines can be tested with a variety of known anticancer agents in order to determine the spectrum of anticancer agents for which the transporter confers resistance; 2)The drug sensitivity of overexpressing cell lines can be used to determine whether newly discovered anticancer agents are transported out of the cell by one of the discovered transporters; 3)Overexpressing cell lines can be used to identify potential inhibitors that reduce the activity of the transporters. Such inhibitors are of great clinical interest in that they may enhance the activity of known anticancer agents, thereby increasing their effectiveness. Reduced activity will be detected by restoration of anticancer drug sensitivity, or by reduction of transporter mediated cellular efflux of anticancer agents. In vitro biochemical studies designed to identify reduced transporter activity in the presence of potential inhibitors can also be performed using membranes prepared from overexpressing cell lines; and 4)Overexpressing cell lines can also be used to

determine whether pharmaceutical agents that are not anticancer agents are transported out of the cell by the transporters.

The following protocols are provided to facilitate the practice of the present invention.

#### Isolation of MOAT-B cDNA

Forward {CT(A/G/T) GT(A/G/T) GC(A/G/T) GT(A/G/T) GT(A/G/T) GG(A/G/C/T)} (SEQ ID NO:9) and reverse {(G/A)CT (A/G/C/T)A(A/G/C) (A/G/C/T)GC (A/G/C/T) (G/C) (T/A) (A/G/C/T)A(A/G) (A/G/C/T)GG (A/G/C/T)TC (A/G)TC} (SEQ ID NO:16) degenerate oligonucleotide primers were designed based upon the first nucleotide binding folds of human MRP, CFTR, and MDR1. Bacteriophage DNA isolated from a C200 cDNA library prepared in the  $\lambda$ pCEV27 phagemid vector (17) was used as template in PCR reactions containing 250 ng cDNA, 5  $\mu$ M primers, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 3 mM MgCl<sub>2</sub>, .05% gelatin, 0.2 mM dNTP and Taq polymerase (Perkin Elmer Cetus). Five cycles of PCR were performed as follows: 94°C for 1 minute, 40°C for 2 minutes, 72°C for 3 minutes. Twenty five cycles were then performed as follows: 94°C for 1 minute, 55°C for 1 minute, and 72°C for 1 minute. The resulting reaction products were used as template in a second round of PCR, as described above, with nested forward {CGGGATCC AG(A/G) GA(A/G) AA(C/T) AT(A/C/T) CT(A/G/C/T) TTT GG(A/G/C/T)} (SEQ ID NO:17) and reverse {CGGAATTC (A/G/T/C)TC (A/G)TC (A/C/T)AG (A/G/C/T)AG (A/G)TA (A/T/G)AT (A/G)TC} (SEQ ID NO:18) degenerate oligonucleotide primers. PCR reaction products were isolated from an agarose gel and subcloned into the BamHI and EcoRI sites of pBluescript (Stratagene). Nucleotide sequence analysis

was performed on plasmid DNA prepared from ampicillin resistant transformants. Additional cDNA clones were isolated from C200 (ovary) and B5 (breast) cDNA libraries by plaque hybridization using the PCR product as the initial radiolabeled probe.

#### RNA Blot Analysis

Blots containing polyA<sup>+</sup> RNA isolated from human tissues (Clontech) were prehybridized at 45°C for 8 hours in 50% formamide, 4X SSC, 4X Denhardt's solution, 0.04 M sodium phosphate monobasic, pH 6.5, 0.8% (w/v) glycine, 0.1 mg/ml sheared denatured salmon sperm DNA. Hybridization was performed at 45°C with <sup>32</sup>P-labeled MOAT-B or GAPDH probes in a solution containing 50% formamide, 3X SSC, 0.04 M sodium phosphate pH 6.5, 10% dextran sulfate, 0.1 mg/ml sheared denatured salmon sperm DNA. Blots were washed 2 times for 15 min at 65°C in 2X SSC, 5 mM Tris-HCl pH 7.4, 0.5% SDS, 2.5 mM EDTA, 0.1% sodium pyrophosphate pH 8.0, and subsequently washed 2 times for 15 min in 0.1X SSC. Blots were then subjected to autoradiography.

#### Chromosomal localization

Preparation of metaphase spreads from phytohemagglutinin-stimulated lymphocytes of a healthy female donor, and fluorescence *in situ* hybridization and detection of immunofluorescence were carried out as previously described (18). A 2.2-kb cDNA clone of MOAT-B inserted in pBluescript was biotinylated by nick translation in a reaction containing 1 µg DNA, 20 µM each of dATP, dCTP and dGTP, 1 µM dTTP, 25 mM Tris-HCl, pH 7.5, 5 mM MgCl<sub>2</sub>, 10 mM β-mercaptoethanol, 10 µM biotin-16-dUTP (Boehringer Mannheim), 2 units DNA polymerase 1/DNase 1 (GIBCO, BRL) and water to a total volume of 50 µl. The



probe was denatured and hybridized to metaphase spreads overnight at 37°C. Hybridization sites were detected with fluorescein-labeled avidin (Oncor) and amplified by addition of anti-avidin antibody (Oncor) and a second layer of fluorescein-labeled avidin. The chromosome preparations were counterstained with DAPI and observed with a Zeiss Axiophot epifluorescence microscope equipped with a cooled charge coupled device camera (Photometrics, Tucson AZ) operated by a Macintosh computer work station. Digitized images of DAPI staining and fluorescein signals were captured, pseudo-colored and merged using Oncor Image version 1.6 software.

#### **Isolation of MOAT-C and MOAT-D cDNA**

MOAT-C and MOAT-D cDNA clones were isolated by plaque hybridization from bacteriophage cDNA libraries using the I.M.A.G.E. clones as the initial probes (ATCC).

#### **RNA blot analysis**

Blots containing polyA<sup>+</sup> RNA isolated from human tissues (Clontech) were purchased from Clontech, and hybridized with radiolabeled MOAT-C, MOAT-D or actin probes according to the manufacturer's directions.

#### **Chromosomal localization**

Preparation of metaphase spreads from phytohemagglutinin-stimulated lymphocytes of a healthy female donor, and fluorescence *in situ* hybridization and detection of immunofluorescence were carried out as previously described (18). A MOAT-C probe inserted in pBluescript, or MOAT-D probe inserted in pBluescript, was biotinylated by nick translation in a reaction containing 1  $\mu$ g DNA, 20  $\mu$ M each of dATP, dCTP and dGTP, 1  $\mu$ M dTTP, 25

mM Tris-HCl, pH 7.5, 5 mM MgCl<sub>2</sub>, 10 mM  $\beta$ -mercaptoethanol, 10  $\mu$ M biotin-16-dUTP (Boehringer Mannheim), 2 units DNA polymerase 1/DNase 1 (GIBCO, BRL) and water to a total volume of 50  $\mu$ l. The probe was denatured and hybridized to metaphase spreads overnight at 37°C. Hybridization sites were detected with fluorescein-labeled avidin (Oncor) and amplified by addition of anti-avidin antibody (Oncor) and a second layer of fluorescein-labeled avidin. The chromosome preparations were counterstained with DAPI and observed with a Zeiss Axiophot epifluorescence microscope equipped with a cooled charge coupled device camera (Photometrics, Tucson AZ) operated by a Macintosh computer work station. Digitized images of DAPI staining and fluorescein signals were captured, pseudo-colored and merged using Oncor Image version 1.6 software.

The following examples are provided to illustrate various embodiments of the invention. They are not intended to limit the invention in any way.

#### EXAMPLE I

##### Isolation of MOAT-B cDNA.

A degenerate PCR approach was used to isolate MRP-related transporters. Degenerate oligonucleotide primers were prepared based upon the N-terminal nucleotide binding folds of MRP and other eukaryotic transporters, and used in conjunction with DNA prepared from an ovarian cancer cell line bacteriophage library. Nucleotide sequence analysis of one of the resulting PCR products indicated that it encoded a segment of a novel nucleotide binding fold that was most closely related to MRP and cMOAT. Overlapping cDNA clones were isolated from ovarian and breast bacteriophage libraries by plaque hybridization using the PCR product as the initial probe. A total of

5.9 kB of cDNA was isolated. Nucleotide sequence analysis revealed two classes of cDNA clones that were about equally represented among isolates from each of the two bacteriophage libraries. The first class contained an open reading frame of 3975 bp that was bordered by in frame stop codons located at positions -76 and -42 (relative to the putative initiation codon) and 3976, and encoding a predicted protein of 1325 amino acids, which is designated MOAT-B. The open reading frame was followed by approximately 2 kB of 3' untranslated sequences. The most upstream ATG in the open reading frame was located in the sequence context "CAAGATGC". The A at position -3 of the putative translation initiation codon was in agreement with the major feature of the Kozak consensus sequence, but the C at position +4 was divergent from the more usual G. The second class of cDNA clones was identical to the first with the exception of a single nucleotide. These clones harbored an additional T following nucleotide 3872 of the first class of clones, close to the C-terminus of the predicted protein. This additional nucleotide resulted in a frame shift such that the predicted protein of the second class of cDNA clones was 22 residues shorter than that of the first class of cDNA clones, and in which the C-terminal 34 residues of the latter reading frame were replaced by 12 distinct residues. See brief description of Figure 1.

#### **Analysis of the MOAT-B Predicted Structure.**

Comparison of the MOAT-B predicted protein with complete coding sequences in protein data bases using the BLAST program indicated that it shared significant similarity with several eukaryotic ABC transporters. Table I.

Table I. Comparison of peptide domains of MOAT-B with those of other eukaryotic ABC transporters

MOAT-B Domain (peptide)	TM1 (88-376)	NBF1 (428-576)	linker region (577-705)	TM2 (706-992)	NBF2 (1058- 1216)	C- terminus (1217- 1325)	overall identity
percent identity							
MRP human	28.6	55.6	27.9	33.3	61.6	51.6	39.2
YCF1 yeast	27	56	27.9	34	57.2	48.5	38.9
MOAT human	33.2	53.3	32.8	31.4	55.3	44.9	38
CFTR Human	30.5	48	27.9	37.7	44	21	36.3
SUR rat	28.1	41.3	28.2	30	52.8	42.8	32.9
MDR1 human	17.6	39.2	21.1	17.3	32.2	40.3	23.3

<sup>B</sup> The indicated domains are, TM1: segment containing the transmembrane spanning domain N-terminal to NBF1; NBF1 and NBF2: nucleotide binding folds 1 and 2; Linker region: segment located between NBF1 and TM2; TM2: segment containing the transmembrane spanning domain located between the two NBFs; C-terminus: segment between NBF2 and the C-terminus of the proteins. Sequence alignments were generated using the PILEUP program of the GCC package. Percent amino acid identity with MOAT-B domains are shown.

Typical features of eukaryotic ABC transporters were present in the predicted MOAT-B protein. See Figure 1. Overall the protein was composed of a tandem repeat of a nucleotide binding fold appended C-terminal to a hydrophobic domain that contained several potential transmembrane spanning helices. Conserved Walker A and B ATP binding sites were present in each of the nucleotide binding folds. See Figure 2A. In addition, a conserved C motif, the signature sequence of ABC transporters, was present in each nucleotide binding fold. Analysis of potential transmembrane motifs using the TMAP program (19) and an input sequence alignment of MOAT-B and MOAT-C, a transporter highly related to MOAT-B<sup>4</sup>, predicted 12 transmembrane helices with 6 transmembrane segments in

each of the two hydrophobic domains. This 6 + 6 configuration of predicted transmembrane helices is in agreement with topological models proposed for MRP and other ABC transporters (20, 21), and is shown in Figure 1. However, alternative predictions of transmembrane segments were obtained using different program parameters or input sequence alignments. For example, when the TMAP program was used with an input sequence alignment consisting of human MRP, rat cMOAT, rat sulfonyl urea receptor (SUR), human cystic fibrosis conductance regulator (CFTR) and human P-glycoprotein, a 6 + 5 configuration was predicted. The only substantial difference between the latter prediction and the structure shown in Figure 1 is that transmembrane segments 9 (829-853) and 10 (855-878) were replaced by a single predicted transmembrane segment spanning amino acids 847 - 875.

Among ABC transporters, the degree of similarity of the nucleotide binding folds is considered to be the best indicator of functional conservation. Comparison of the nucleotide binding folds of MOAT-B with other eukaryotic ABC transporters indicated that it was most closely related to MRP, the yeast cadmium resistance protein (YCF1) and cMOAT (Table I), three transporters that have organic anions as substrates. The MOAT-B NBF1 was 55.6, 56.0 and 53.3 percent identical, and the MOAT-B NBF2 was 61.6, 57.2 and 55.3 percent identical to the first and second nucleotide binding folds of human MRP, YCF1 and human cMOAT, respectively. Aside from the latter transporters, the MOAT-B nucleotide binding folds were most closely related to those of CFTR and SUR. The MOAT-B nucleotide binding folds shared significantly less similarity with those of MDR1. Alignment of the MOAT-B nucleotide binding folds with those of other eukaryotic

transporters is shown in Figure 2A. Analysis of the overall amino acid identity of MOAT-B with other ABC transporters also indicated that it was most closely related to MRP, YCF1 and cMOAT (Table I). Overall MOAT-B was 39.2, 38.9 and 38 percent identical to these transporters, respectively. Figure 2B shows a comparison of the hydropathy profiles of MOAT-B with those of other eukaryotic transporters. This comparison reveals that MOAT-B (1325 amino acids) is approximately 200 amino acids smaller than MRP (1531 residues), cMOAT (1545 residues) and YCF1 (1515 residues), and that this size difference is largely accounted for by the absence in MOAT-B of an amino terminal hydrophobic extension that is present in MRP, cMOAT and YCF1 (22). This N-terminal hydrophobic segment is predicted to harbor several transmembrane spanning segments, and is also present in SUR.

#### **Expression Pattern of MOAT-B in Human Tissues.**

To gain insight into the possible function of MOAT-B, its expression pattern in a variety of human tissues was examined by RNA blot analysis. As shown in Figure 3, a MOAT-B transcript of approximately 6 kB was readily detected. The isolation of 5.9 kB of MOAT-B cDNA was consistent with this size. MOAT-B expression was detected in each of the 16 tissues analyzed. Transcript levels were highest in prostate and lowest in liver and peripheral blood leukocytes, for which prolonged exposure of film were required to detect expression. Intermediate levels of expression were observed in other tissues.

#### **Chromosomal Localization of the MOAT-B Gene.**

The MOAT-B chromosomal localization was determined by fluorescence *in situ* hybridization. As shown in Figure 4, hybridization of the MOAT-B probe to metaphase spreads revealed specific labeling at human chromosome band 13q32.

Fluorescent signals were detected on chromosome 13 in each of 19 metaphase spreads scored. Of 135 signals observed, 62 (46%) were on 13q. Among these signals, 61 localized at 13q32, near the boundary between 13q31 and 13q32. Paired (on sister chromatids) signals were only seen at band 13q32. In several metaphases, signals on a single chromatid were observed at chromosome bands 6p21 or 4q21, suggesting hybridization to distantly related sequences.

## EXAMPLE II

### Isolation of MOAT-C and MOAT-D cDNA.

Isolation of the MOAT-B<sub>4</sub> transporter as described above suggested the possibility that there were other MRP/cMOAT-related transporters. A blast search (36) of the nonredundant expressed sequence tag data base using MRP and related yeast transporters revealed two clones with significant similarity to MRP and cMOAT. The first of these sequences (I.M.A.G.E. consortium clone 113196) was 1.2 kb in length, 800 bp of which encoded an MRP-related peptide. A segment of this clone was used as a probe to screen ovarian and hematopoietic bacteriophage libraries. Analysis of these cDNA clones indicated that they contained approximately 2 kb of additional coding sequence not present in clone 113196. An additional 1655 bp of 5' sequence was obtained by several rounds of RACE using the bacteriophage DNA prepared from the ovarian cDNA library as template. The continuity of the sequences obtained by RACE with the cDNA clones isolated from bacteriophage libraries was confirmed by nucleotide sequence analysis of a 2 kb product obtained by RT/PCR using an upstream oligonucleotide primer located at the 5' end of the RACE sequence and a downstream primer located at the 5' end of the cDNA obtained by plaque

hybridization. A total of approximately 5.9 kb of cDNA sequences were isolated. Nucleotide sequence analysis revealed an open reading frame of 4311 bp that was preceded by an in frame stop codon located at positions -93 (relative to the putative initiation codon), and encoding a predicted protein of 1437 amino acids, which is designated MOAT-C herein. The open reading frame was followed by approximately 1.4 kb of 3' untranslated sequences in which a polyadenylation sequence (AAUAAA) was located 20 bp upstream of the poly(A) tail. The most upstream ATG in the open reading frame was located in the sequence context "GAAGATGA". The A at position -3 of the putative translation initiation codon was in agreement with the major feature of the Kozak consensus sequence, but the A at position +4 was divergent from the more usual G (37). The second sequence identified in our data base search (I.M.A.G.E. consortium clone 208097) was 1.2 kb in length, of which 588 bp encoded an MRP-related peptide. A segment of this clone was used as a probe to screen liver and monocyte bacteriophage cDNA libraries, and 5' cDNA segments of the isolated cDNA clones were used in a subsequent round of screening. Together approximately 5.2 kb of cDNA sequence were isolated. Nucleotide sequence analysis revealed an open reading frame of 4570 bp, which is designated MOAT-D herein. The open reading frame was followed by approximately 0.6 kb of 3' untranslated sequences in which a polyadenylation sequence (AAUAAA) was located 12 bp upstream of the poly(A) tail. An upstream in frame stop codon was not present in the MOAT-D cDNA clones, and attempts to obtain additional upstream sequences by RACE using as template cDNA prepared from sources in which MOAT-D is abundant were not successful. The most upstream ATG in the open reading frame



(nucleotide position 5-7), located in the sequence context "ATGGATGG", was therefore designated as the translational initiation site. The G at position +4, was in good agreement with the Kozak consensus sequence, but the T at -3 was divergent from the more usual A (37). Although an upstream in frame stop codon was not identified in the MOAT-D cDNA clones, the size of the encoded protein was within one amino acid of the size of the transporter with which it shares the highest degree of identity (MRP), suggesting that the complete MOAT-D open reading frame was present in the isolated cDNA clones.

#### **Analysis of the MOAT-C and MOAT-D Predicted Proteins.**

Comparison of the MOAT-C and MOAT-D predicted proteins with complete coding sequences in protein data bases using the BLAST program indicated that they shared significant similarity with several eukaryotic ABC transporters. Typical features of eukaryotic ABC transporters were present in the predicted proteins. See Figure 5. Overall the proteins were composed of hydrophobic domains containing potential transmembrane spanning helices and two nucleotide binding folds. Conserved Walker A and B ATP binding sites, as well as a conserved C motif, the signature sequence of ABC transporters, was present in the nucleotide binding folds. Computer assisted analysis of potential transmembrane helices of MOAT-C using the TMAP program (19) predicted 12 transmembrane helices with 6 transmembrane spanning helices in each of two membrane spanning domains. This 6 + 6 (TM1-TM6 and TM7-TM12) configuration of predicted transmembrane helices is in agreement with topological models proposed for several other ABC transporters (20, 21), and is shown in Figure 5. However, alternative

predictions of transmembrane segments were obtained using different program parameters or input sequence alignments. Comparison of the hydropathy profiles of MOAT-C with other MRP/cMOAT-related transporters (Fig. 6B) indicates that its structure is similar to that of MOAT-B, which also has two membrane spanning domains.

In contrast to MOAT-C, hydrophobicity analysis of MOAT-D indicated that it has three membrane spanning domains. Similar to MRP, cMOAT and the yeast cadmium resistance factor 1 (YCF1), MOAT-D has an additional N-terminal hydrophobic domain that is not present in MOAT-B or MOAT-C (Figs. 5 and 6). A 5+6+6 configuration of transmembrane spanning helices has been proposed for MRP (38), in which the N-terminal extension harbors 5 transmembrane spanning helices, and 6 transmembrane helices are present in the second and third membrane spanning domain. An alignment of the MOAT-D predicted protein with MRP using the GAP program indicated that proposed MRP transmembrane spanning helices were conserved in MOAT-D. This 5+6+6 model for MOAT-D is shown in Fig. 5. Another configuration of transmembrane spanning helices (5+6+4) was predicted using computer assisted analysis. MRP has been reported to have two N-linked glycosylation sites in its N-terminus (Asn-19 and Asn-23) and another site located between the first and second transmembrane spanning helix of its third membrane spanning domain (Asn-1006). The alignment of MOAT-D with MRP indicated that an N-terminal (Asn-21) and a distal N-glycosylation sites (Asn-1008/1009) were conserved in analogous positions in MOAT-D. Only the distal N-glycosylation site of MRP is conserved in MOAT-C (Asn890) (Fig. 5) and MOAT-B (Asn746/754).

Among ABC transporters, the degree of similarity of

the nucleotide binding folds is considered to be the best indicator of functional conservation. Comparison of the nucleotide binding folds of MOAT-C and MOAT-D with other eukaryotic ABC transporters indicated that they were most closely related to those of human MRP, human cMOAT and yeast YCF1, three transporters that have organic anions as substrates. As shown in Table 2, among the human transporters, the MOAT-C NBF1 was about equally related to MOAT-D, MRP and cMOAT (55-61% identity), and less similar to MOAT-B (49% identity).

Table II. Amino acid identity: nucleotide binding folds 1 and 2 of MRP/cMOAT sub-family members.

	MOAT-C	MOAT-D	MOAT-B	MRP	cMOAT	YCF1
	%IDENTIFY (BNF1/NBF20)					
MOAT-C	-----	57.3/58.9	49.3/59.1	60.0/59.4	61.3/60.6	55.3/58/8
MOAT-D	57.3/58/9	-----	55.3/54.1	70.173.8	67.3/70.0	52.7/61.3
MOAT-B	49.3/59.1	55.3/54.1	-----	57.3/61/6	53.3/55.3	56.0/57.2
MRP	60.0/59.4	70.7/73.7	57.3/61.6	-----	66.0/73.1	53.3/63.8
cMOAT	61/3/60.6	67.3/70.0	53.3/55.3	66.0/73.1	-----	50.7/61/3
YCF1	55.3/58.8	52.7/61.3	56.0/57.2	53.3/63.8	50.7/61.3	-----

The MOAT-C NBF2 shared about equal amino acid identity with the five other transporters in this group (59-61% identity). Overall, the MOAT-C protein was about equally related to the other five transporters in this group, with 33.1-36.5% identity. Aside from these

transporters, MOAT-C is most closely related to CFTR, with which its NBFs shared 44%/42 % identity, and SUR, with which its NBFs shared 49%/51% identity.

The MOAT-D NBFs were clearly most closely related to those of MRP and cMOAT, with which they shared considerable amino acid identity (67.3-73.8%). See Table III. Of the latter two transporters, the MOAT-D NBFs were slightly more related to those of MRP. In contrast, the MOAT-D NBFs shared only 55.3-58.9% identity with those of MOAT-C and MOAT-B. Overall, MOAT-D was again most closely related to MRP (57.3%) and cMOAT (46.9%), but significantly more related to MRP. Consistent with the analysis of NBFs, MOAT-D was much less related to MOAT-C and MOAT-B, with which it shared only 33.1% and 35.3% identity, respectively. Alignment of the MOAT-C and MOAT-D nucleotide binding folds with those of other eukaryotic transporters is shown in Fig. 6.

Table III. Overall amino acid identifying among MRP/cMOAT sub-family members

	MOAT-C	MOAT-D	MOAT-B	MRP	cMOAT	YCF1
	%identity					
MOAT-C	----	33.1	36.5	35.8	36.2	33.6
MOAT-D	33.1	----	35.3	57.3	46.9	38.1
MOAT-B	36.4	35.3	----	39.4	36.8	38.8
MRP	35.8	57.3	39.4	----	48.4	46.4
cMOAT	36.3	46.9	36.8	48.8	----	38.8
YCF1	33.6	38.1	38.8	40.4	38.8	----

#### **Expression Pattern of MOAT-C and MOAT-D in Human Tissues.**

To gain insight into the possible functions of MOAT-C and MOAT-D, their expression patterns in a variety of human tissues was examined by RNA blot analysis. As

shown in Fig. 7 (upper panels), a MOAT-C transcript of approximately 6.6 kB was readily detected in several tissues. MOAT-C transcript levels were highest in skeletal muscle, with intermediate levels in kidney, testes, heart and brain. Low levels were detected in most other tissues, including spleen, thymus, prostate, ovary, and placenta. Prolonged exposures were required for detection in lung and liver. MOAT-D was expressed as an approximately 6 kb transcript (middle panels). Compared to MOAT-C, the MOAT-D expression pattern was more restricted. MOAT-D was highly expressed in colon and pancreas, with lower levels in liver and kidney. Low levels were detected in small intestine, placenta and prostate. Prolonged exposures were required to detect MOAT-D in testes, thymus, spleen and lung.

#### **Chromosomal localization of the *MOAT-C* and *MOAT-D* genes.**

The *MOAT-C* and *MOAT-D* chromosomal localizations were determined by fluorescence *in situ* hybridization. As shown in Figure 8, hybridization of the MOAT-C probe to metaphase spreads revealed specific labeling at human chromosome band 3q27. Fluorescent signals were detected on chromosome 3q in each of 22 metaphase spreads scored. Of 75 signals observed, 43 (57%) were on 3q. Paired (on sister chromatids) signals were only seen at band 3q27. Hybridization of the MOAT-D probe revealed specific labeling at human chromosome band 17q21.3. Fluorescent signals were detected on chromosome 17 in each of 21 metaphase spreads scored. Of 83 signals observed, 34 (41%) were on 17q21.3. Paired (on sister chromatids) signals were only seen at band 17q21.3.

### EXAMPLE III

#### Isolation of MOAT-E and MOAT-E cDNA.

Analysis of ara, a reported cDNA sequence that encodes a 453 amino acid transporter, revealed that it is a non-physiological sequence representing a combination of 5' MRP sequences fused to an MRP/cMOAT-related transporter. The MRP sequences extend to codon 8 of the reported predicted protein.

To isolate the complete physiological cDNA, a RT/PCR approach was employed in which primers were designed based upon a reported genomic sequence that encodes exons identical to the reported ara sequence. The MOAT-E cDNA was isolated in three segments. The first segment, spanning residues 1-616, was isolated by PCR using 5' primer ATGGCCGCGCCTGCTGAGC; (SEQ ID NO: 10) and 3' primer GTCTACGACACCAGGGTCAA (SEQ ID NO: 11). The second segment, spanning residues 1815-3187, was isolated by PCR using 5' CTGCCTGGAAGAAGTTGACC (SEQ ID NO: 12) and 3' primer CTGGAATGTCCACGTCAACC (SEQ ID NO: 13). The third segment, spanning residues 3158-1503, was isolated by PCR using 5' primer GGAGACAGACACGGTTGACG (SEQ ID NO: 14) and 3' primer GCAGACCAGGCCTGACTCC (SEQ ID NO: 15). The primer were designed based upon the nucleotide sequence of human genomic BAC clone CIT987SD-962B4. The template for these reactions was random-primed human kidney cDNA prepared from total RNA. Using this approach the physiological cDNA was isolated which is designated MOAT-E herein and set forth as Sequence I.D. No. 7.

#### Analysis of the MOAT-E Predicted Protein.

MOAT-E encodes a 1503 amino acid transporter. The MOAT-E predicted amino acid sequence is designated Sequence I.D. No. 8. See Figure 9. Also shown is the

location of potential transmembrane helices (overbars), potential N-glycosylation site (black dot) and the two nucleotide binding folds (NBF1 and NBF2). Walker A and B motifs, as well as the signature C motif of ABC transporters are also indicated. Comparison of MOAT-E with ara indicates that the ara predicted protein is not only a fused sequence, but also that it represents only 446 (~30%) of the 1503 MOAT-E residues.

Comparison of MOAT-E with the other members of the MRP/cMOAT subfamily, which include MRP, cMOAT, MOAT-B, MOAT-C and MOAT-E, is shown in Table IV. MOAT-E is highly related to MOAT-D, MRP and cMOAT, with which it shares 39-45% identity. This high degree of identity is also indicated by the high percent identities of the nucleotide binding folds, which range from 55-61%. In contrast, MOAT-E is less related to MOAT-B and MOAT-C, with which it shares ~31% and 34% identity, respectively.

Table IV. Amino acid identity among MRP/cMOAT sub-family members.<sup>a</sup> The bold type indicates the percent identity of the overall proteins, and the parentheses indicates the percent identity of the nucleotide binding folds.

	MOAT-E	MOAT-B	MOAT-C	MOAT-D	MRP	cMOAT
	% identity <sup>b</sup>					
MOAT-E	---	<b>33.9</b>	<b>30.6</b>	<b>43.6</b>	<b>45.1</b>	<b>38.9</b>
	---	(52.0/56.6)	(50.0/52.5)	(59.3/59.4)	(61.3/61.4)	(55.3/59.4)
MOAT-B	<b>33.9</b>	---	<b>36.4</b>	<b>35.3</b>	<b>39.4</b>	<b>36.8</b>
	(52.0/56.6)	---	(49.3/59.1)	(55.3/54.1)	(57.3/61.6)	(56.0/57.2)
MOAT-C	<b>30.0</b>	<b>36.4</b>	---	<b>33.1</b>	<b>35.8</b>	<b>36.2</b>
	(50.0/52.5)	(49.3/59.1)	---	(57.3/58.9)	(60.6/59.4)	(61.3/60.6)
MOAT-D	<b>43.6</b>	<b>35.3</b>	<b>33.1</b>	---	<b>57.3</b>	<b>46.9</b>
	(59.3/59.4)	(55.3/54.1)	(57.3/58.9)	---	(70.7/73.8)	(67.3/70.0)
MRP	<b>45.1</b>	<b>39.4</b>	<b>35.8</b>	<b>57.3</b>	---	<b>48.4</b>
	(61.3/61.9)	(57.3/61.6)	(60.0/59.4)	(70.7/73.8)	---	(66.0/73.1)
cMOAT	<b>38.9</b>	<b>36.8</b>	<b>36.2</b>	<b>46.9</b>	<b>48.4</b>	---
	(53.1/59.4)	(56.0/57.2)	(61.3/60.6)	(67.3/70.0)	(66.0/73.1)	---

<sup>a</sup>overall amino acid identities are indicated in bold-face, and identities of nucleotide binding folds 1 and 2 are indicated in parentheses (NBF1/NBF2).

<sup>b</sup>percent identity was obtained using the GAP command in the GCG package.

Comparison of the hydropathy profile of MOAT-E with other members of the MRP/cMOAT subfamily is shown in figure 10. The data reveal that MOAT-E has a hydrophobic N-terminal segment that is present in its closest relatives, MOAT-D, MRP and cMOAT. This structural feature is present in all of the currently known organic anion transporters, and suggests that MOAT-E may share substrate specificity with MRP and cMOAT. MOAT-E may also share the drug resistance activity of the latter two proteins. In contrast, MOAT-B and MOAT-C do not have this hydrophobic N-terminal extension.

#### **Expression Pattern of MOAT-E in Human Tissues.**

In a Northern blot of RNA isolated from various tissues, MOAT-E expression is restricted to liver and kidney, suggesting that MOAT-E may participate the excretion of substances into the urine and bile. See Figure 11. This figure also shows that MOAT-E is expressed as an ~6 kB transcript. This is in contrast to the ~2.3 kB transcript that was reported for ara, clearly indicating that the fused ara transcript is unique to the cell line from which it was isolated, and is not a physiological transcript. Together, the isolation of MOAT-E and analysis of its sequence and expression pattern suggest that it may be involved in cellular resistance to drugs and/or the excretion of drugs into the urine and bile.

#### **DISCUSSION**

The present invention discloses additional MRP/cMOAT-related transporters which were identified by



using a degenerative PCR cloning approach in which the conserved amino terminal ATP-binding domain of known eukaryotic transporters was targeted. Using this approach the complete coding sequences of MOAT-B, MOAT-C, MOAT-D and MOAT-E were obtained. MOAT-B is a protein whose predicted structure indicates that it is a member of the ABC transporter family. Comparison of the MOAT-B predicted protein with other transporters reveals that it is most closely related to MRP, cMOAT and yeast YCF1, and thus extends the number of known full length MRP-related transporters. The similarity of MOAT-B to these transporters suggest that it shares a similar substrate specificity. Transport assays using membrane vesicle preparations indicate that MRP is capable of transporting diverse organic anions, including glutathione S-conjugates such as LTC<sub>4</sub>, oxidized glutathione, and glucuronidated and sulfated conjugates of steroid hormones and bile salts (7). Although membrane vesicle transport assays of substrate specificity using cMOAT-transfected cells have not yet been reported, genetic and biochemical studies using TR- and EHBR rat strains, which are defective in the hepatobiliary excretion of glutathione and glucuronate conjugates, indicate that it is also an ATP-dependent transporter of organic anions. cMOAT, which is primarily expressed in the canalicular membrane of hepatocytes, has been reported to be absent in these rat strains, and hepatocyte canalicular membranes prepared from the mutant rats are deficient in the ATP-dependent transport of glutathione and glucuronate conjugates (23, 24). In addition, cMOAT protein has also been reported to be absent in the hepatocytes of patients with Dubin-Johnson syndrome (25), a disorder manifested by chronic

conjugated hyperbilirubinemia. YCF1, a yeast transporter, has also been demonstrated to transport glutathione complexes (26). Thus, based upon the similarity of MOAT-B to these three transporters, it is possible that it also functions to transport organic anions, an activity critical to the cellular detoxification of a wide range of xenobiotics.

MOAT-C, MOAT-D and MOAT-E are three other MRP/cMOAT-related transporters. The isolation of these two transporters extends the number of known full length members of this subfamily to six. Based upon the degree of amino acid similarity and overall topology these six proteins fall into two groups. The first group is composed of MOAT-D, MOAT-E, MRP and cMOAT. These four transporters are highly related, sharing ~39-45% amino acid identity. MOAT-D is more closely related to MRP (57% identity) than is cMOAT (48% identity), and is therefore the closest known relative of MRP. In addition to a high degree of amino acid identity, the similarity between MOAT-D, MRP and cMOAT, also extends to overall topology. Like MRP and cMOAT, MOAT-D and MOAT-E have three membrane spanning domains, including an N-terminal hydrophobic extension that is predicted to harbor ~5 transmembrane helices, and which is absent in transporters such as CFTR and MDR1. This N-terminal extension is also present in YCF1, a related yeast transporter that transports glutathione S-conjugates, and SUR, a more distantly related transporter involved in the regulation of potassium channels. The second group of MRP/cMOAT-related transporters is composed of MOAT-B and MOAT-C. These two transporters are distinguished from the first group by their lower level of amino acid similarity and distinct topology. Like MOAT-D and MOAT-E, MOAT-B

and MOAT-C are more closely related to MRP (39% and 36%, respectively) and cMOAT (37% and 36%, respectively) than to other eukaryotic transporters. However, they share considerably less similarity with MRP, cMOAT, MOAT-D and MOAT-E than the latter four transporters share with each other (~39-45% identity). In addition, in contrast to MRP, cMOAT, MOAT-D and MOAT-E, MOAT-B and MOAT-C do not have an N-terminal membrane spanning domain, and their topology is therefore more similar to many other eukaryotic ABC transporters that also have only two membrane spanning domains.

Defining the contributions of MOAT-B, MOAT-C, MOAT-D and MOAT-E to cytotoxic drug resistance will facilitate the design of novel chemotherapeutic agents. The multidrug resistance activity of MRP is well described. While the drug sensitivity pattern of cMOAT-transfected cells has not yet been reported, the possibility that it may also confer resistance to cytotoxic drugs is suggested by a recent report in which transfection of a cMOAT antisense vector was found to enhance the sensitivity of a human liver cancer cell line to both natural product drugs and cisplatin. Since MOAT-D and MOAT-E are more closely related to MRP than is cMOAT, the possibility that they will also confer resistance is particularly intriguing. The availability of the MOAT-B, MOAT-C, MOAT-D and MOAT-E cDNAs will facilitate the analysis of their possible contributions to cytotoxic resistance.

## References

1. Gottesman, M. M. and Pastan, I. Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annu. Rev. Biochem.*, 62: 385-427, 1993
2. Kruh, G. D., Chan, A., Myers, K., Gaughan, K., Miki, T., and Aaronson, S. A. Expression complementary DNA library transfer establishes mrp as a multidrug resistance gene. *Cancer Res.*, 54: 1649-52, 1994.
3. Zaman, G. J., Flens, M. J., van Leusden, M. R., de Haas, M., Mulder, H.S., Lankelma, J., Pinedo, H. M., Scheper, R. J., Baas, F., Broxterman, H.J., and Borst, P. The human multidrug resistance-associated protein MRP is a plasma membrane drug-efflux pump. *Proc. Natl. Acad. Sci. U S A*, 91: 8822-6, 1994.
4. Grant, C. E., Valdimarsson, G., Hipfner, D. R., Almquist, K. C., Cole, S.P., and Deeley, R. G. Overexpression of multidrug resistance-associated protein (MRP) increases resistance to natural product drugs. *Cancer Res.*, 54:357-61, 1994.
5. Breuninger, L. M., Paul, S., Gaughan, K., Miki, T., Chan, A., Aaronson, S. A., and Kruh, G. D. Expression of Multidrug Resistance-associated Protein in NIH/3T3 Cells Confers Multidrug Resistance Associated with Increased Drug Efflux and Altered Intracellular Drug Distribution. *Cancer Res.*, 55: 5342-5347, 1995.
6. Cole, S. P., Sparks, K. E., Fraser, K., Loe, D. W., Grant, C. E., Wilson, G. M., and Deeley, R. G. Pharmacological characterization of multidrug resistant MRP-transfected human tumor cells. *Cancer Res.*, 54: 5902-10, 1994.
7. Keppler, D., Leier, I., and Jedlitschky, G. Transport of glutathione conjugates and glucuronides by the multidrug resistance proteins MRP1 and MRP2. *Biol. Chem.*, 378: 787-91, 1997.
8. Lee, J. S., Scala, S., Matsumoto, Y., Dickstein, B., Robey, R., Zhan, Z., Altenberg, G., and Bates, S. E. Reduced drug accumulation and multidrug resistance in human breast cancer cells without associated P-glycoprotein or MRP overexpression. *J. Cell. Biochem.*, 65: 513-26, 1997.
9. Gately, D. P. and Howell, S. B. Cellular accumulation

of the anticancer agent cisplatin: a review. Br. J. Cancer, 67: 1171-6, 1993.

10. Ishikawa, T. and Ali-Osman, F.

Glutathione-associated cis-diamminedichloroplatinum(II) metabolism and ATP-dependent efflux from leukemia cells. Molecular characterization of glutathione-platinum complex and its biological significance. J. Biol. Chem., 268: 20116-25, 1993.

11. Ishikawa, T., Wright, C. D., and Ishizuka, H. GS-X pump is functionally overexpressed in cis-diamminedichloroplatinum (II)-resistant human leukemia HL-60 cells and down-regulated by cell differentiation. J. Biol. Chem., 269: 29085-93, 1994.

12. Fujii, R., Mutoh, M., Sumizawa, T., Chen, Z. S., Yoshimura, A., and Akiyama, S. Adenosine triphosphate-dependent transport of leukotriene C4 by membrane vesicles prepared from cisplatin-resistant human epidermoid carcinoma tumor cells [see comments]. J. Natl. Cancer Inst., 86: 1781-4, 1994.

13. Ishikawa, T., Bao, J. J., Yamane, Y., Akimaru, K., Frindrich, K., Wright, C. D., and Kuo, M. T. Coordinated induction of MRP/GS-X pump and gamma-glutamylcysteine synthetase by heavy metals in human leukemia cells. J. Biol. Chem., 271: 14981-8, 1996.

14. Taniguchi, K., Wada, M., Kohno, K., Nakamura, T., Kawabe, T., Kawakami, M., Kagotani, K., Okumura, K., Akiyama, S., and Kuwano, M. A human canalicular multispecific organic anion transporter (cMOAT) gene is overexpressed in cisplatin-resistant human cancer cell lines with decreased drug accumulation. Cancer Res., 56: 4124-9, 1996.

15. Kool, M., de Haas, M., Scheffer, G. L., Scheper, R. J., van Eijk, M. J., Juijn, J. A., Baas, F., and Borst, P. Analysis of expression of cMOAT (MRP2), MRP3, MRP4, and MRP5, homologues of the multidrug resistance-associated protein gene (MRP1), in human cancer cell lines. Cancer Res., 57: 3537-47, 1997.

16. Koike, K., Kawabe, T., Tanaka, T., Toh, S., Uchiumi, T., Wada, M., Akiyama, S., Ono, M., and Kuwano, M. A canalicular multispecific organic anion transporter (cMOAT) antisense cDNA enhances drug sensitivity in human hepatic cancer cells. Cancer Res., 57: 5475-9, 1997.

17. Miki, T., Fleming, T. P., Crescenzi, M., Molloy, C. J., Blam, S. B., Reynolds, S. H., and Aaronson, S. A. Development of a highly efficient expression cDNA cloning system: application to oncogene isolation. *Proc. Natl. Acad. Sci. U S A*, 88: 5167-71, 1991.
18. Bell, D. W., Taguchi, T., Jenkins, N. A., Gilbert, D. J., Copeland, N. G., Gilks, C. B., Zweidler-McKay, P., Grimes, H. L., Tsichlis, P. N., and Testa, J. R. Chromosomal localization of a gene, GF1, encoding a novel zinc finger protein reveals a new syntenic region between man and rodents. *Cytogenet. Cell. Genet.*, 70: 263-7, 1995.
19. Persson, B. and Argos, P. Prediction of transmembrane segments in proteins utilising multiple sequence alignments. *J. Mol. Biol.*, 237: 182-92, 1994.
20. Bakos, E., Hegedus, T., Hollo, Z., Welker, E., Tusnady, G. E., Zaman, G. J., Flens, M. J., Varadi, A., and Sarkadi, B. Membrane topology and glycosylation of the human multidrug resistance-associated protein. *J. Biol. Chem.*, 271: 12322-6, 1996.
21. Loo, T. W. and Clarke, D. M. Membrane topology of a cysteine-less mutant of human P-glycoprotein. *J. Biol. Chem.*, 270: 843-8, 1995.
22. Tusnady, G. E., Bakos, E., Varadi, A., and Sarkadi, B. Membrane topology distinguishes a subfamily of the ATP-binding cassette (ABC) transporters. *FEBS Lett.*, 402: 1-3, 1997.
23. Paulusma, C. C., Bosma, P. J., Zaman, G. J., Bakker, C. T., Otter, M., Scheffer, G. L., Scheper, R. J., Borst, P., and Oude Elferink, R. P. Congenital jaundice in rats with a mutation in a multidrug resistance-associated protein gene. *Science*, 271: 1126-82, 1996.
24. Buchler, M., Konig, J., Brom, M., Kartenbeck, J., Spring, H., Horie, T., and Keppler, D. cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance protein, cMrp, reveals a novel conjugate export pump deficient in hyperbilirubinemic mutant rats. *J. Biol. Chem.*, 271: 15091-8, 1996.
25. Kartenbeck, J., Leuschner, U., Mayer, R., and Keppler, D. Absence of the canalicular isoform of the MRP gene-encoded conjugate export pump from the hepatocytes

- in Dubin-Johnson syndrome. *Hepatology*, 23: 1061-6, 1996.
26. Li, Z. S., Szczypka, M., Lu, Y. P., Thiele, D. J., and Rea, P. A. The yeast cadmium factor protein (YCF1) is a vacuolar glutathione S-conjugate pump. *J. Biol. Chem.*, 271: 6509-17, 1996.
27. Wemmie, J. A. and Moye-Rowley, W. S. Mutational analysis of the *Saccharomyces cerevisiae* ATP-binding cassette transporter protein Ycf1p. *Mol. Microbiol.*, 25: 683-94, 1997.
28. Kruh, G. D., Gaughan, K. T., Godwin, A. K., and Chan, A. Expression Pattern of MRP in Human Tissues and Adult Solid Tumor Cell Lines. *J. Natl. Cancer Inst.*, 87: 1256-58, 1995.
29. Longhurst, T. J., O'Neill, G. M., Harvie, R. M., and Davey, R. A. The anthracycline resistance-associated (ara) gene, a novel gene associated with multidrug resistance in a human leukaemia cell line. *Br. J. Cancer*, 74: 1331-5, 1996.
30. Allikmets, R., Gerrard, B., Hutchinson, A., and Dean, M. Characterization of the human ABC superfamily: isolation and mapping of 21 new genes using the expressed sequence tags database. *Hum. Mol. Genet.*, 5: 1649-55, 1996.
31. Shen, D.-w., Pastan, I., and Gottesman, M. M. Cross-Resistance to Methotrexate and Metals in Human Cisplatin-resistant Cell Lines Results from a Pleiotropic Defect in Accumulation of These Compounds Associated with Reduced Plasma Membrane Binding Proteins. *Cancer Res.*, 58: 268-275, 1998.
32. Naredi, P., Heath, D. D., Enns, R. E., and Howell, S. B. Cross-resistance between cisplatin and antimony in a human ovarian carcinoma cell line. *Cancer Res.*, 54: 6464-8, 1994.
33. Naredi, P., Heath, D. D., Enns, R. E., and Howell, S. B. Cross-resistance between cisplatin, antimony potassium tartrate, and arsenite in human tumor cells. *J. Clin. Invest.*, 95: 1193-8, 1995.
34. Wemmie, J. A., Szczypka, M. S., Thiele, D. J., and Moye-Rowley, W. S. Cadmium tolerance mediated by the yeast AP-1 protein requires the presence of an

ATP-binding cassette transporter-encoding gene, YCF1. J. Biol. Chem., 269: 32592-7, 1994.

35. O'Dwyer, P. J., Johnson, S. W., and Hamilton, T. C. Cisplatin and its Analogues. In: V. T. J. DeVita, S. Hellman, and S. A. Rosenberg (eds.), Cancer Principles and Practice of Oncology, pp. 418-432. Philadelphia: Lippincott-Raven, 1997.

36. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D. J. Basic Local Alignment Search Tool. J. Mol. Biol. 215:403-10, 1990.

37. Kozak, M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nuc. Acids. Res. 12:857-72.

38. Tusnady, G.E., Bakos, E., Varadi, A., Sarkadi, B. Membrane topology distinguishes a subfamily of the ATP-binding cassette (ABC) transporters. FEBS Lett. 402:1-3, 1997.

While certain of the preferred embodiments of the present invention have been described and specifically exemplified above, it is not intended that the invention be limited to such embodiments. Various modifications may be made thereto without departing from the scope and spirit of the present invention, as set forth in the following claims.



What is claimed is:

1. An isolated nucleic acid molecule having the sequence of SEQ ID NO:1, said nucleic acid molecule comprising a nucleotide sequence encoding a MOAT-B transporter protein about 1350 amino acids in length, said encoded transporter protein comprising a multi-domain structure including a tandem repeat of nucleotide binding folds appended C-terminal to a hydrophobic domain, said nucleotide binding folds having Walker A and B ATP binding sites, said C-terminal domain having a plurality of membrane spanning helices.

2. The nucleic acid molecule of claim 1, which is DNA.

3. The DNA molecule of claim 2, which is a cDNA comprising a sequence approximately 5.9 kilobase pairs in length that encodes said MOAT-B transporter protein.

4. The DNA molecule of claim 2, which is a gene comprising introns and exons, the exons of said gene specifically hybridizing with the nucleic acid of SEQ ID NO 1, and said exons encoding said MOAT-B transporter protein.

5. An isolated RNA molecule transcribed from the nucleic acid of claim 1.

6. The nucleic acid molecule of claim 1, wherein said sequence encodes a MOAT-B transporter

protein having an amino acid sequence selected from the group consisting of SEQ ID NO 2 and amino acid sequences encoded by natural allelic variants of said sequence.

7. The nucleic acid molecule of claim 6, which comprises SEQ ID NO 1.

8. An antibody immunologically specific for the protein encoded by the nucleic acid of claim 1.

9. An antibody as claimed in claim 8, said antibody being monoclonal.

10. An antibody as claimed in claim 8, said antibody being polyclonal.

11. An isolated nucleic acid molecule having the sequence of SEQ ID NO: 3, said nucleic acid molecule comprising a sequence encoding a MOAT-C transporter protein about 1450 amino acids in length, said transporter protein having a multi-domain structure including a tandem repeat of nucleotide binding folds, said nucleotide binding folds having Walker A and B binding sites, and a C-terminal hydrophobic domain that contains several membrane spanning helices.

12. The nucleic acid molecule of claim 11, which is DNA.

13. The DNA molecule of claim 12, which is a cDNA comprising a sequence approximately 6.6 kilobase pairs in length that encodes said MOAT-C transporter protein.

14. The DNA molecule of claim 12, which is a gene comprising introns and exons, the exons of said gene specifically hybridizing with the nucleic acid of SEQ ID NO 3, and said exons encoding said MOAT-C transporter protein.

15. An isolated RNA molecule transcribed from the nucleic acid of claim 11.

16. The nucleic acid molecule of claim 11, wherein said sequence encodes a MOAT-C transporter protein having an amino acid sequence selected from the group consisting of SEQ ID NO 4 and amino acid sequences encoded by natural allelic variants of said sequence.

17. The nucleic acid molecule of claim 11, which comprises SEQ ID NO 3.

18. An antibody immunologically specific for the protein encoded by the nucleic acid of claim 11.

19. An antibody as claimed in claim 18, said antibody being monoclonal.

20. An antibody as claimed in claim 18, said antibody being polyclonal.

21. An oligonucleotide between about 10 and about 200 nucleotides in length, which specifically hybridizes with a protein translation initiation site in a nucleotide sequence encoding amino acids of SEQ ID NO 4.

22. An oligonucleotide between about 10 and about 200 nucleotides in length, which specifically hybridizes with a protein translation initiation site in a nucleotide sequence encoding amino acids of SEQ ID NO 2.

23. An isolated nucleic acid molecule having the sequence of SEQ ID NO: 5, said nucleic acid molecule comprising a sequence encoding a MOAT-D transporter protein about 1550 amino acids in length, said transporter protein having a multi-domain structure including a tandem repeat of nucleotide binding folds, said nucleotide binding folds having Walker A and B binding sites, and a C-terminal hydrophobic domain that contains several membrane spanning helices.

24. The nucleic acid molecule of claim 23, which is DNA.

25. The DNA molecule of claim 24, which is a cDNA comprising a sequence approximately 6 kilobase pairs in length that encodes said MOAT-D transporter protein.

26. The DNA molecule of claim 24, which is a gene comprising introns and exons, the exons of said gene specifically hybridizing with the nucleic acid of SEQ ID NO 5, and said exons encoding said MOAT-D transporter protein.

27. An isolated RNA molecule transcribed from the nucleic acid of claim 23.

28. The nucleic acid molecule of claim 23, wherein

said sequence encodes a MOAT-D transporter protein having an amino acid sequence selected from the group consisting of SEQ ID NO 6 and amino acid sequences encoded by natural allelic variants of said sequence.

29. The nucleic acid molecule of claim 23, which comprises SEQ ID NO 5.

30. An antibody immunologically specific for the protein encoded by the nucleic acid of claim 23.

31. An antibody as claimed in claim 30, said antibody being monoclonal.

32. An antibody as claimed in claim 30, said antibody being polyclonal.

33. An oligonucleotide between about 10 and about 200 nucleotides in length, which specifically hybridizes with a protein translation initiation site in a nucleotide sequence encoding amino acids of SEQ ID NO 6.

34. An isolated nucleic acid molecule having the sequence of SEQ ID NO:7, said nucleic acid molecule comprising a nucleotide sequence encoding a MOAT-E transporter protein about 1503 amino acids in length, said transporter protein having a multi-domain structure including a tandem repeat of nucleotide binding folds, said nucleotide binding folds having Walker A and B binding sites, and a C-terminal hydrophobic domain that contains several membrane spanning helices.

35. The nucleic acid molecule of claim 34,

which is DNA.

36. The DNA molecule of claim 35, which is a cDNA comprising a sequence approximately 6 kilobase pairs in length that encodes said MOAT-E transporter protein.

37. The DNA molecule of claim 35, which is a gene comprising introns and exons, the exons of said gene specifically hybridizing with the nucleic acid of SEQ ID NO 7, and said exons encoding said MOAT-E transporter protein.

38. An isolated RNA molecule transcribed from the nucleic acid of claim 34.

39. The nucleic acid molecule of claim 34, wherein said sequence encodes a MOAT-E transporter protein having an amino acid sequence selected from the group consisting of SEQ ID NO 8 and amino acid sequences encoded by natural allelic variants of said sequence.

40. The nucleic acid molecule of claim 39, which comprises SEQ ID NO 7.

41. An antibody immunologically specific for the protein encoded by the nucleic acid of claim 34.

42. An antibody as claimed in claim 41, said antibody being monoclonal.

43. An antibody as claimed in claim 41, said antibody being polyclonal.

44. An oligonucleotide between about 10 and about 200 nucleotides in length, which specifically hybridizes with a protein translation initiation site in a nucleotide sequence encoding amino acids of SEQ ID NO 7.

45. A plasmid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 and SEQ ID NO:7.

46. A vector comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 and SEQ ID NO:7.

47. A retroviral vector comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 and SEQ ID NO:7.

48. A host cell comprising at least one nucleic acid molecule having a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 and SEQ ID NO:7.

49. A host cell as claimed in claim 48, wherein said host cell is selected from the group consisting of bacterial, fungal, mammalian, insect and plant cells.

50. A host cell as claimed in claim 48, wherein said nucleic acid is provided in a plasmid and is operably linked to mammalian regulatory elements which confer high expression and stability of mRNA transcribed from said nucleic acid.

51. A host cell as claimed in claim 48, wherein said nucleic acid is provided in a plasmid and is operably linked to mammalian regulatory control elements in reverse anti-sense orientation.

52. A host animal comprising at least one nucleic acid molecule selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 and SEQ ID NO: 7.

53. A host animal as claimed in claim 52, wherein said animal harbors a homozygous null mutation in its endogenous MOAT gene wherein said mutation has been introduced into said mouse or an ancestor of said mouse via homologous recombination in embryonic stem cells, and further wherein said mouse does not express a functional mouse MOAT protein.

54. The transgenic mouse of claim 53, wherein said mouse is fertile and transmits said null mutation to its offspring.

55. The transgenic mouse of claim 53, wherein said null mutation has been introduced into an ancestor of said mouse at an embryonic stage following microinjection of embryonic stem cells into a mouse blastocyt.

56. A method for screening a test compound for inhibition of MOAT mediated transport, comprising:

a) providing a host cell expressing at least one MOAT-encoding nucleic acid having a sequence selected from the group consisting of SEQ ID NOS: 1, 3, 5, and 7;



b) contacting said host cell with a compound suspected of inhibiting MOAT-mediated transporter activity; and

c) assessing inhibition of transport mediated by said compound.

57. A method as claimed in claim 56, wherein inhibition of MOAT mediated transport is indicated by restoration of anticancer drug sensitivity.

58. A method as claimed in claim 57, wherein said inhibition of MOAT mediated transport is indicated by a reduction of transporter mediated cellular efflux of anticancer agents.

59. A kit for detecting the presence of MOAT encoding nucleic acids in a sample, comprising:

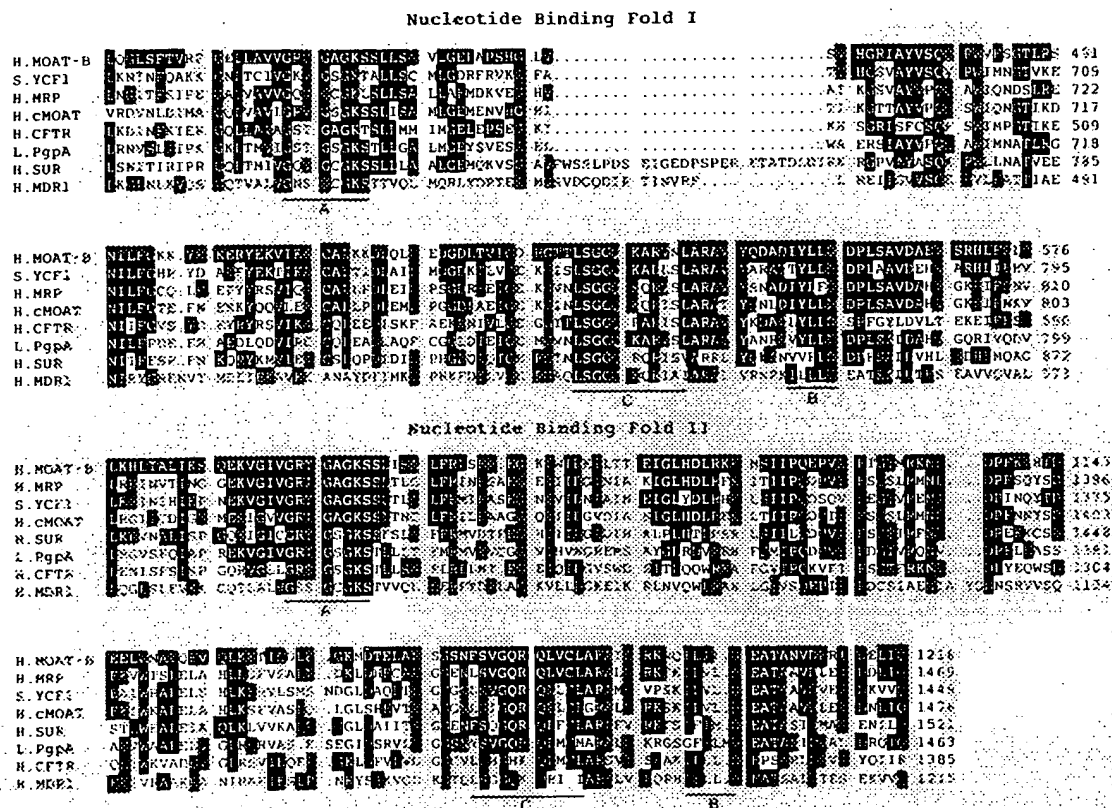
- a) oligonucleotide primers specific for amplification of MOAT encoding nucleic acids;
- b) polymerase enzyme;
- c) amplification buffer; and
- d) MOAT specific DNA for use as a positive control.

1/56

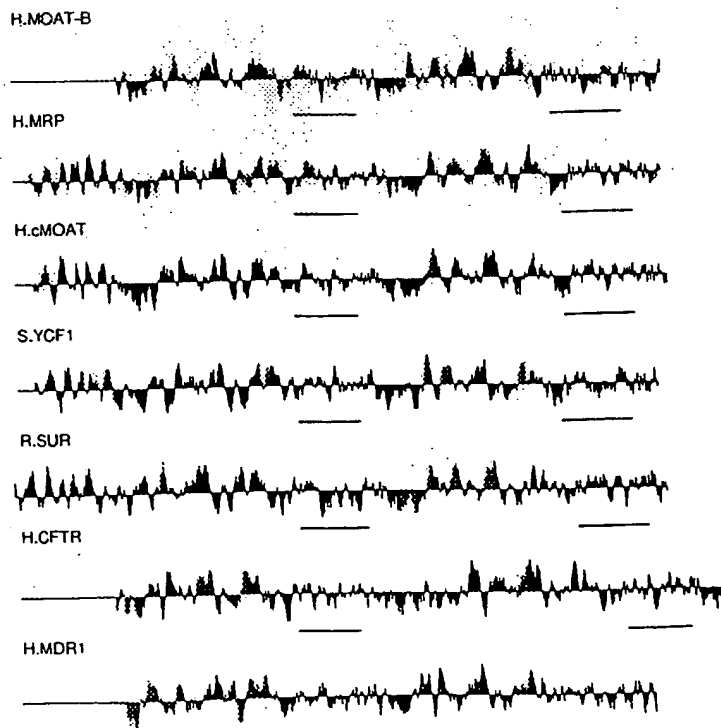
MOAT-B .....  
MRP 1 MALRGFCSADGSDPLWDMVNTWNTSNPDFTKCFQNTVLVWVPCFYLVACFPFYFLYLSRHDRGYIQMTPLNKTALGFLHIVCWADLFYSFWERSRGI 100  
MOAT-B 1 .....MLP 3  
MRP 101 FLAPVFLVSPITLLGITTLLATFLIQLERRKGVQSSGIMLTFLWLVALCALAILRSKIMTALKEDAQVDLFRDITFVYVYFSLLIQLVLSCFSDRSPLFSE 200  
MOAT-B 4 VYQEVKPNPLQDANICSRVFFWMLNPLFKIGHKRRLEEDDMYSVLPEDRSOHLGEELQGFWDKEVLRAENDAOK ..... 77  
MRP 201 TIHDPNPCESSASFSLRITFWMITGLIVRGYRQPLEGSDLSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQPKESSKVDANEVEAL 300  
MOAT-B 78 .....PSLTRAIIKCYWKSylvLGIFTLIEESAKVIOPIFLGKIINYFENVDPHDSVALNTAYAYATVLTFTLILAILHHLYFYHVQCAGHRL 166  
MRP 301 IVKSPOKEWNPFLFKVLYKTFGPYFLMSFFFKATHDLHMFSGPQILKLLIKFVNDTKAPDWQGY.....FYTVLLFVTACLQTLVLHQYFHICFVSGMRI 395  
MOAT-B 167 RVAMCHMIYRKALRLSNMAMGKTTTGQIVNLLSNDVNKFDQVTVFLHFLWAGPLQAI AVTALLWMEIGISCLAGHAVLIILLPLQSCFGKLFSSLSRSKTA 266  
MRP 396 KTAVIGAVYRKALVITNSARKSSTVGEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLWNLGPSVLGAVVMVHLVVPVNAVMAHMKTKTYQVAHM 495  
MOAT-B 267 TFTDARIITHNEVITGIRIIMKYAWEKSFNSLITNLRKKEISKILRSSCLRGQNLASFFSASKIIVFVTFTTYVLLG...SVITASRVFVAVTLYGAVRLT 364  
MRP 496 KSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQEELKVLKKSAYLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDAQTAFLVSLALFNILRFP 595  
MOAT-B 365 VTLEFFPSAIERVSEAIIVSIRRIQTFLLEIS...QRNRQLPSDGKMHVHQDFTAFWDKASETPTLQGLSFTVRPCELLAVVGPVGAGKSSLLSAVLG 460  
MRP 596 LNI.LPMVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGNTSITVRNATFTWAR.SDPPTLNGITFSIPEGALVAVVGQVCGKLSLLSALLA 693  
MOAT-B 461 ELAPSHGLVSVHGRIAYVSQQPWVFSCTLRSLNIFGKKYKERYEKVIKACALKKDLQLEDGLTVIGDRGTTLSGGQKARVNLAARAVYQDADIYLLDD 560  
MRP 694 EMDKVEGHVAIKGSVAVYVQQAQWQNDLRENILFGCCOLEEPPYRSVIOACALLPDLEILPSGDRTEIGEGVNLSCGQKQVRLARAVYSNADIYLLDD 793  
MOAT-B 561 PLSAVDAEVSRLHFLCICQ...ILHEKITILVTHQQLYKLAASQILILKDGKMHVQKGTYTEFLKSGIDFGSLK...KDNESSEQPPVPG... 645  
MRP 794 PLSAVDAHVGHKIFENVIGPKMLKNTIRILVTHMSYLPQVDVIVHSGGKISEMGVQELLARDGAFAELRTYASTEQQEQAENGVTCVSGPGKEA 893  
MOAT-B 646 .....TPTLRNRTFSESSVMSQSSRPSLKDGALESQDT...ENVPTLSEENRSEKGVGFQAYKNYFRAGAHWIVFIFLILLNTAAQVAVYVQ 731  
MRP 894 KQENGHLVTDGAKQLQRLSSSSSYSGDISRHNNSTAELQKAEAKKEETWKLMEADKAQTCQVKLSVYWDYHKAICLFIISFLIFL.MCNHVSALAS 992  
MOAT-B 732 DWVLSYWANKQSHLNTVNGCGNVTEKLDLWYLGITYSGLTVATVLFGIARSLVYFVYVNLSSQTLHNKMFESILKAPVLFDRNPGRILNRFKSDIGH 831  
MRP 993 NYWLSLWTD...DPIVNGTQEHKVR...LSVYGALGISQCAVFGYSHAVSIGGILASRCLHVDLLHSILRSPHSFFERTPSGNLVNRFKSKELDT 1082  
MOAT-B 832 LDDLPLTFLDFIOTLLQVGVVSVAVAVIPWIAIPLVPLGIIIFIILRRYFLETSRDVVKLESTTRSPVFSHLSSSLOGLWITIRAYKAERQELFDAHQ 931  
MRP 1083 VDSMIEVIMKFMGSLFNIVIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSRQLKRLSESRSVPYSHFNETLLGVSVIRAFEEQERFIHQSDLVK 1182  
MOAT-B 932 DLHSEAWFLFTTSRWFAVRLDAICAMFVIIAFGSLILAKTLDAQVGLALSYALTLMGHFQWCVROSAEVENMMISVERVIEYTDLEKEAPWEYQK.R 1030  
MRP 1183 DENQKAYYPSIVANRWLAVRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVVRMSSEHETNIVAVERLKEYSETEKEAPWQIQETR 1282  
MOAT-B 1031 PPPAWPHEGVIIIFDNVNFHYSFGGPLVKHLTALIKSOEKVIGVGRTGAGKSSLSALFRLSE.PEGKIWDKILTTEICGLHDLRKHSIIPOEPVLFTG 1129  
MRP 1283 PPSSHPQVGRVEFRNYCLRYREDLDFVLRHINVTINGGEKVGIVGRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSG 1382  
MOAT-B 1130 THRNKLDPPFKEHTDEELWNLQEVOLKETIEDLPKMDTELAESGSNFSVGQRQLVCLARAILRKNQIILIDEATANVDPRTDELIOKKIREKFAHCTVL 1229  
MRP 1383 SLRNLDPPFSQVSDDEEVWTSLELAHLKDFVSALPKDLOHECAECGENLSVGQRQLVCLARALLRKTKILVLDATAAVDLETODLIQSTIRTQFEDCTVL 1482  
MOAT-B 1230 TIAHRLNTIIDSCKIMVLDSCRLKEYDEPVYLLONKESLFYKMHVQOLGKAEAAALTETAKQVYFKRNYHIGHTDHKMTNTSNGOPSTLTIFETAL 1325  
MRP 1483 TIAHRLNTIMDYTRVIVLDKGEIQEYCAPSDLLQOR.GLFYSMAKDAGLV 1531

Figure 1

**Fig. 2A**



**Fig. 2B**



3/56

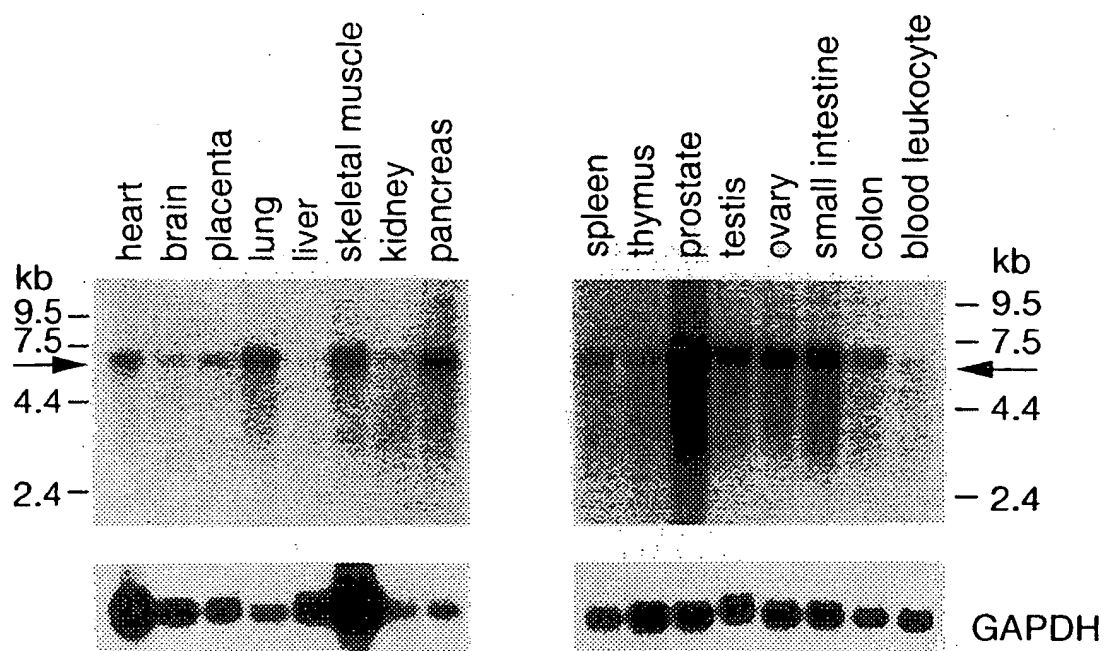


Figure 3

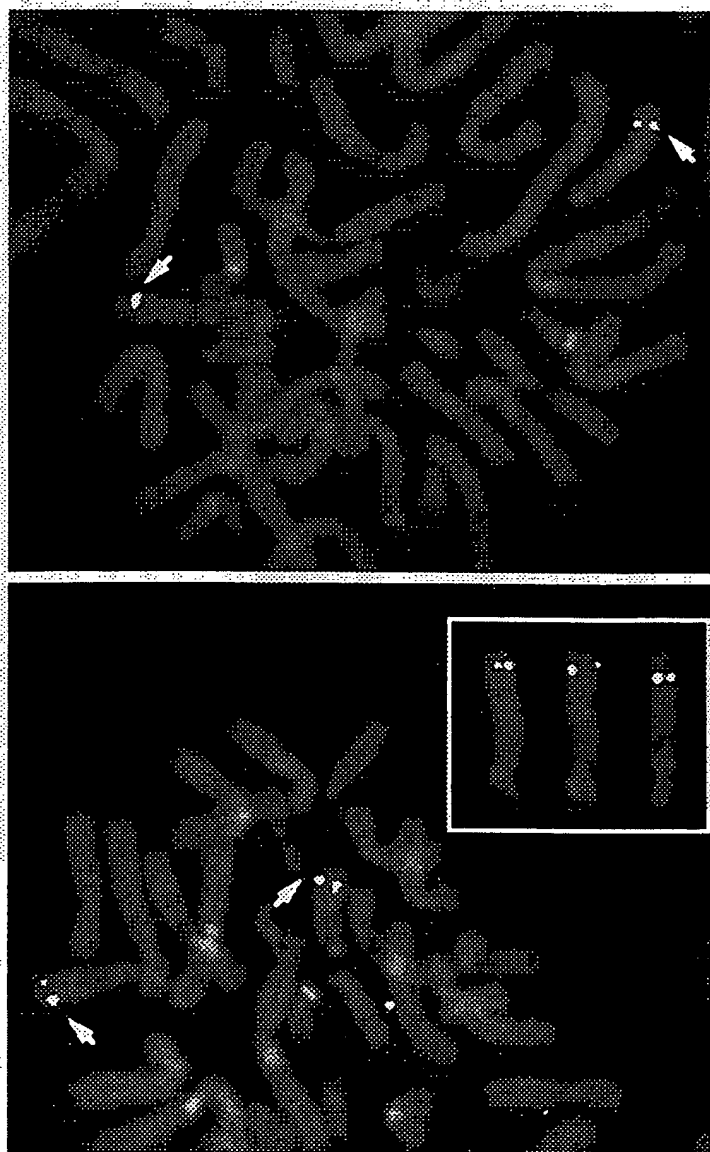


Figure 4

5/56

Fig. 5A

1 MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFRR TRPLECQDAL ETAARAEGLS  
 61 LDASHMSQLR ILDEEHPK GK YHHGLSALKP IRTTSKHQHP VDNAGLFSM TFSWLSLAR  
 121 VAHKKGELSM EDVWSLSKHE SSDVNCRRLE RLWQEELNEV GPDAASLRV VWIFCRTL  
     TM1 TM2  
 181 LSIVCLMITQ LAGFSGPAPM VKHLEYTQA TESNLQYSL LVLGLLLEI VRSWSLALTW  
     TM3  
 241 ALNYRTGVRL RGAILTMAFK KILKLKNIKE KSLGELINIC SNDGQRMFEA AAVGSLLAGG  
     TM4  
 301 PVVAILGMIY NVIILGPTGF LGSVAVFILFY PAMMFASRLT AYFRRCVAA TDERVQKMNE  
     TM5  
 361 VLTYIKFIKM YAWVKAFSQS VQKIREEERR ILEKAGYFQG ITVGVAPIVV VIASVVTFSV  
     TM6  
 421 HMTLGFDLTA AQAPT VVTVF NSMTFALKVT PFSVKSLSEA SVAVDRFKSL FLMEEVHMIK  
 481 NKPASPHIKI EMKNATLAWD SSHSSIQNSP KLT PKMKKDK RASRGKKEKV RQLQRTHEQA  
     NBF1  
 541 VLAEQKGHLL LDSDERPSPE EEEGKHIHLG HRLQRTLS IDLEIQEGKL VGICGSVSGG  
     A  
 601 KTSLSAILG QMTLLEGSIA ISGTFAIVAQ QAWILNATLR DNILFGKEYD EERYNSVLNS  
 661 CCLRPDLAIL PSSDLTEIGE RGANLSCGQR QRISLARALY SDRSIYILDD PLSALDAHVG  
     NBF1 C B  
 721 NHIFNSAIRK HLKSKTVLFV THQLQYLVD C DEVIFMKEGC ITERGTHEEL MNLNGDYATI  
 781 FNNLLGETP PVEINSKKT SGSQKKSQDK GPKTGSVKKE KAVKPEEGQL VQLEEKQGS  
     TM7  
 841 VPWSVYGVYI QAAGGPLAFL VIMALFMLNV GSTAFSTWWL SYWIKQSGN TTVTRGNETS  
     TM8  
 901 VDSMKDNPH MQYYASIAL SHAVMLILKA IRGVVFVKGT LRASSRLHDE LFRRILRSPM  
     TM9  
 961 KFFDTTPTGR ILNRFSDMD EVDVRLPFQA EMFIQNVILV FFCVGMIAGV FFWFLVAVGP  
     TM10  
 1021 LVILFSVLHI VSRVLIRELK RLDNITQSPF LSHITSSIOG LATIHAYNKG QEFLHRYQEL  
     TM11 TM12  
 1081 LDDNQAPFFL FTCAMRWLAV RDLISIALI TTTGLMTVLM HGQIPPAYAG LAISTAVQLT  
 1141 GLFQFTVRLA SETEARFTSV ERINHYIKL SLEAPARIKN KAPSPDWPQE GEVTFENAEM  
     NBF2  
 1201 RYRENPLVL KVSFTIKPK EKIGIVGRTG SGKSSLGMAL FRLVELSGGC IKIDGVRISD  
     A  
 1261 IGLADLRSL SIIPQEPVLF SGTVRNLD PNOYTEDQIW DALERTHME CIAQLPLKE  
     NBF2  
 1321 SEVMENGDNF SVGERQLLCI ARALLRHCKI LILDEATAAM DTETOLLIQE TIREAFADCT  
     C B  
 1381 MLTIAHRLHT VLGS DRIMVL AQGQVVEFDT PSVLLSNDSS RFYAMFAAAE NKVAVKG

6/56

Fig. 5B

1 MGPM DALCGS GELGSKFWD S NLSVHTENPD LTPCFQNSLL AWWPCYILWV ALPCYLLYL R TM1  
 61 HHCRGYIILS HLSKLMVLG VLLWCVSWAD LFYSFHGLVH GRAPAPVFFV TPLVVGVTML TM2  
 121 LATLLIQYER LQGVQSSGVL IIFWFLCVVC AIVPFRSKIL LAKAEGEISD PFRFTTFYIH TM3 TM4  
 181 FALVLSALIL ACFREKPPFF SAKNVDPNPY PETS VGFLSR LFFWWFTKMA IYGYRHPLEE TM5  
 241 KDLWSLKEED RSQMVVQQLL EAWRKQEKQT ARHKASAAPG KNASGEDEV L LGARPRPRK TM6  
 301 SFLKALLATF GSSFLISACF KLIQDLLSFI NPQLLSILIR FISNPMAPSW WGFLVAGLMF TM7  
 361 LCSMMQSLIL QHYHYIFVT GVKFRTGIMG VIYRKALVIT NSVKRASTVG EIVNLM SVDA TM8  
 421 QRFMDLAPFL NLLWSAPLOI ILAIYFLWQN LGPSVLAGVA FMVLLIPLNG AVAVKMR AFQ TM9  
 481 VKQMKLDSR IKLMSEILNG IKVLKLYAWE PSFLKQVEGI RQGELOLLRT AAYLHTTTT TM10  
 541 TWM CSPFLVT LITLWVYVYV DPNNVLDAEK AFVSVSLFNI LRLPLNMLPQ LISNLTQASV TM11  
 601 SLKRIQQFLS QEELDQSV E RTISPGYAI TIHSGTFTWA QDL PPTLHSL DIQVPKGALV NBF1  
 661 AVVGPVGC GK SSLVSALLGE MEKLEGKVHM KGSVAYVPOQ AWIQNCTLOE NVLF GKALNP A  
 721 KRYQQTLEAC ALLADLEMLP GGDQTEIGEK GINLSGGORQ RVSLARAVYS DADIFLLDDP NBF1 C B  
 781 LSAVD SHVAK HIFDHVIGPE GVLAGKTRVL VTEGISFLPQ TDFIIVLADG QVSEMGPIPA  
 841 LLQRNGSFAN FLCNYAPDED QGHLEDSWTA LEGAEDKEAL LIEDTLSNHT DLTNDNPVTY  
 901 VVQKQFMRQL SALSSDGEQ GRPVPRRH LG PSEKVQVTEA KADGALTQEE KAAIGTVELS TM12  
 961 VFWDYAKAVG LCTTLAICLL YVGQSA AAG ANVWLSAWTN DAMADSRQNN TSLRLGVYAA TM13  
 1021 LGILQGFLVM LAAMAMAAGG IQAARVLEQA LLHNKIRSPQ SFFDTTPSGR ILNCF SKDIY TM14 TM15  
 1081 VVDEV LAPVI LMLLNSFFNA ISTLVVIMAS TPLFTVVILP LAVLYTLVQR FYAATSRQLK  
 1141 RLESVSRSPI YSHFSETVTG ASVIRAYNRS RDEFIISDTK V DANQRSCYP YIISNRWLSI TM16 TM17  
 1201 GVEFVGNCVV LFAALFAVIG RSSLNPGLVG LSVSYSLQVT FALNWMIRM SDLESNIVAV NBF2  
 1261 ERVKEYSKTE TEAPWVVEGS RPPEGWPPRG EVEFRNYSVR YRPGLDLVLR DLSLHVHGGE  
 1321 KVGIVGRTGA GKSSMTLC LF RILEAAKGEI RIDGLNVADI GLHDLRSQLT IIPQDPILFS A  
 1381 GTLRMNLDPF GSYSEEDIWV ALELSHLETF VSSQAGLDF OCSEGGENLS VGORQLVCLA NBF2 C  
 1441 RALLRKSRI L VLDEATAAID LETDNLIQAT IRTOFDTCTV LTIAHRLNTI MDYTRVLVLD B  
 1501 KGVVAEFDSP ANLIAARGIF YGMARDAGLA

7/56

## Nucleotide Binding Fold I

MOAT-D	SHSLDTQVPK	GALVAVVGPV	GGKSSIVSA	LUREVKIEG	KI	.....	.....	HM	KESVAYVPC	ATPCSCLCG	707
MRP	INGTFSTPE	GALVAVVGPV	GGKSSLLSA	LLAEVQKVEG	HI	.....	.....	AI	KESVAYVPC	AMJONDSLRE	724
cMOAT	VRDVNLDMA	GALVAVVGPV	GGKSSLLSA	MUSEVENVIG	HI	.....	.....	TI	KETAYVPC	STICGFIKD	717
MOAT-C	IHSIDLEQE	GKLVGICSS	SGHTSIISA	IGQMTLIEG	SI	.....	.....	AI	SGTFAYVPC	ATLAILLRD	641
MOAT-B	IGGLSFTVRP	GELVAVVGPV	GGKSSLLSA	VGGELAPSHG	LI	.....	.....	SV	HRIAYVPC	PAVFSGLRS	491
CFTR	IKDNFKER	GGLLAIAST	PAKTSLLMM	IMGLEPSEC	KI	.....	.....	KH	SGRISFCS	FSMPGILKE	504
SUR	LSNFTIRPR	SGLTHIKQK	GGKSSLLSA	ALAEVQKVEG	AFWSSLPDS	EIGEDPSPER	ETATOLDIRK	RSPVAYVPC	PALLIAVVE	785	
MDR1	IKGLNLKVQS	SGTVALKNS	GGKSSITVOL	MORLYDPTES	MSVDCQDIR	TINVRFLREI	.....	IGVVEE	PVLFATIAE	486	

A

MOAT-D	NVLFKA.LN	PKRYQOTLEA	CALLADDEMD	EGGLOFEIGE	KGINLGGQOR	QFVSLARAVY	SEADIFELDD	PLSAVDSSVA	KHIDHV	793
MRP	NILECCO.LE	EPYVRSVICA	CALLEDEHEI	EGGLOFEIGE	KGINLGGQOR	QFVSLARAVY	SNADITLDD	PLSAVDHNG	KHIDENV	810
cMOAT	NILEFTE.FN	SKRYOCVLEA	CALLEDEHEI	EGGLOFEIGE	KGINLGGQOR	QFVSLARAVY	ONLDIYLLDD	PLSAVDHNG	KHIDENV	803
MOAT-C	NILEFKE.YD	SEYNSVLNS	CGRPDLAI	EGGLOFEIGE	KGINLGGQOR	QFVSLARAVY	SRSTIYLLDD	PLSALDANG	KHIDNSA	727
MOAT-B	NILEKK.YE	KRYEYVICA	CALKKDLQLE	EDGGLVIGD	EGTFLSGGQR	ARVNLARAVY	QWADIYLLDL	PLSAVDSSVA	KHIDENV	577
CFTR	NILEVS.YD	EPYVRSVICA	CGEELDISK	AEKINIVLSE	GGTFLSGGQR	ARVNLARAVY	KWADIYLLDL	PLSAVDSSVA	KHIDENV	590
SUR	NILESP.FN	KRYEYVICA	CGEELDISK	AEKINIVLSE	GGTFLSGGQR	ARVNLARAVY	QWADIYLLDL	PLSAVDSSVA	KHIDENV	871
MDR1	IRYRENVNT	MDEIEKAVKE	ANAYDFIMK	HKFDLVVEG	EGGLOFEIGE	KGINLGGQOR	QFVSLARAVY	SEADIFELDD	PLSAVDSSVA	573

C

B

## Nucleotide Binding Fold II

MOAT-D	IRDLHLVNG	CEKVGVIGRT	JACKSSMTLC	LEFILAANK	ERREGLNVR	DGLHDLRSQ	RTIPCEHIL	FSGILMML	DPFGSTEE	1392
MRP	IRHINVTNG	CEKVGVIGRT	JACKSSMTLC	LEFILAANK	ERREGLNVR	DGLHDLRSQ	RTIPCEHIL	FSGILMML	DPFGSTEE	1396
cMOAT	IRGITCDSS	CEKVGVIGRT	JACKSSMTLC	LEFILAANK	ERREGLNVR	DGLHDLRSQ	RTIPCEHIL	FSGILMML	DPFGSTEE	1403
MOAT-C	IKKVEFTAP	KEFVIGVGR	GGKSSLMGA	LEFILAANK	ERREGLNVR	DGLHDLRSQ	RTIPCEHIL	FSGILMML	DPFGSTEE	1296
MOAT-B	IKHLTALIKS	CEKVGVIGRT	JACKSSMTLC	LEFILAANK	ERREGLNVR	DGLHDLRSQ	RTIPCEHIL	FSGILMML	DPFGSTEE	1143
SUR	IKHVNALSP	KEFVIGVGR	GGKSSLMGA	LEFILAANK	ERREGLNVR	DGLHDLRSQ	RTIPCEHIL	FSGILMML	DPFGSTEE	1447
CFTR	ENISFSTSP	KEFVIGVGR	GGKSSLMGA	LEFILAANK	ERREGLNVR	DGLHDLRSQ	RTIPCEHIL	FSGILMML	DPFGSTEE	1312
MDR1	OGSILEVKK	QOTLALVSS	CEKSTVVQL	LEFILAANK	ERREGLNVR	DGLHDLRSQ	RTIPCEHIL	FSGILMML	DPFGSTEE	1142

A

MOAT-D	EDITWALELS	NLHTVSSCF	AGLDFOCES	GENLSVGGR	IVCLARALLR	RSRIIVLDEA	TAADILETN	LIR	1465
MRP	EDITWALELS	NLHTVSSCF	AGLDFOCES	GENLSVGGR	IVCLARALLR	RSRIIVLDEA	TAADILETN	LIR	1469
cMOAT	EDITWALELS	NLHTVSSCF	AGLDFOCES	GENLSVGGR	IVCLARALLR	RSRIIVLDEA	TAADILETN	LIR	1476
MOAT-C	EDITWALELS	NLHTVSSCF	AGLDFOCES	GENLSVGGR	IVCLARALLR	RSRIIVLDEA	TAADILETN	LIR	1369
MOAT-B	EDITWALELS	NLHTVSSCF	AGLDFOCES	GENLSVGGR	IVCLARALLR	RSRIIVLDEA	TAADILETN	LIR	1216
SUR	EDITWALELS	NLHTVSSCF	AGLDFOCES	GENLSVGGR	IVCLARALLR	RSRIIVLDEA	TAADILETN	LIR	1385
CFTR	EDITWALELS	NLHTVSSCF	AGLDFOCES	GENLSVGGR	IVCLARALLR	RSRIIVLDEA	TAADILETN	LIR	1520
MDR1	EDITWALELS	NLHTVSSCF	AGLDFOCES	GENLSVGGR	IVCLARALLR	RSRIIVLDEA	TAADILETN	LIR	1215

C

B

Fig. 6A



8/56

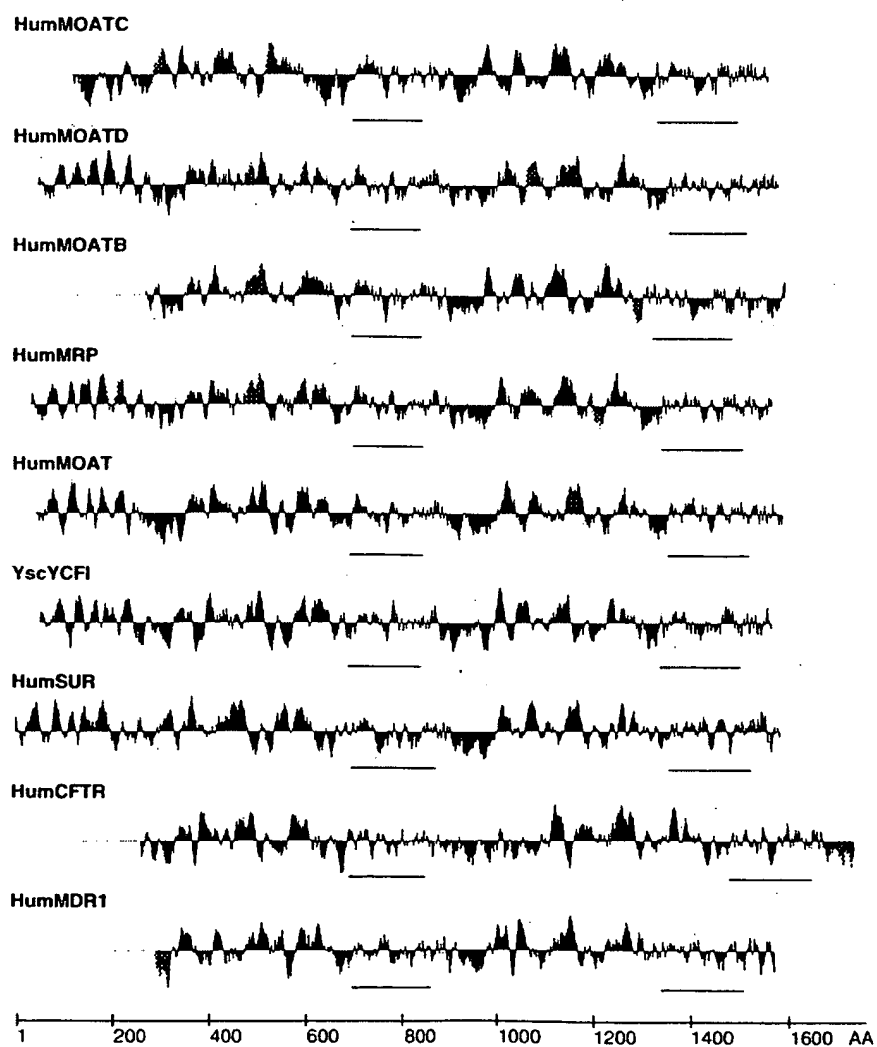


Fig. 6B

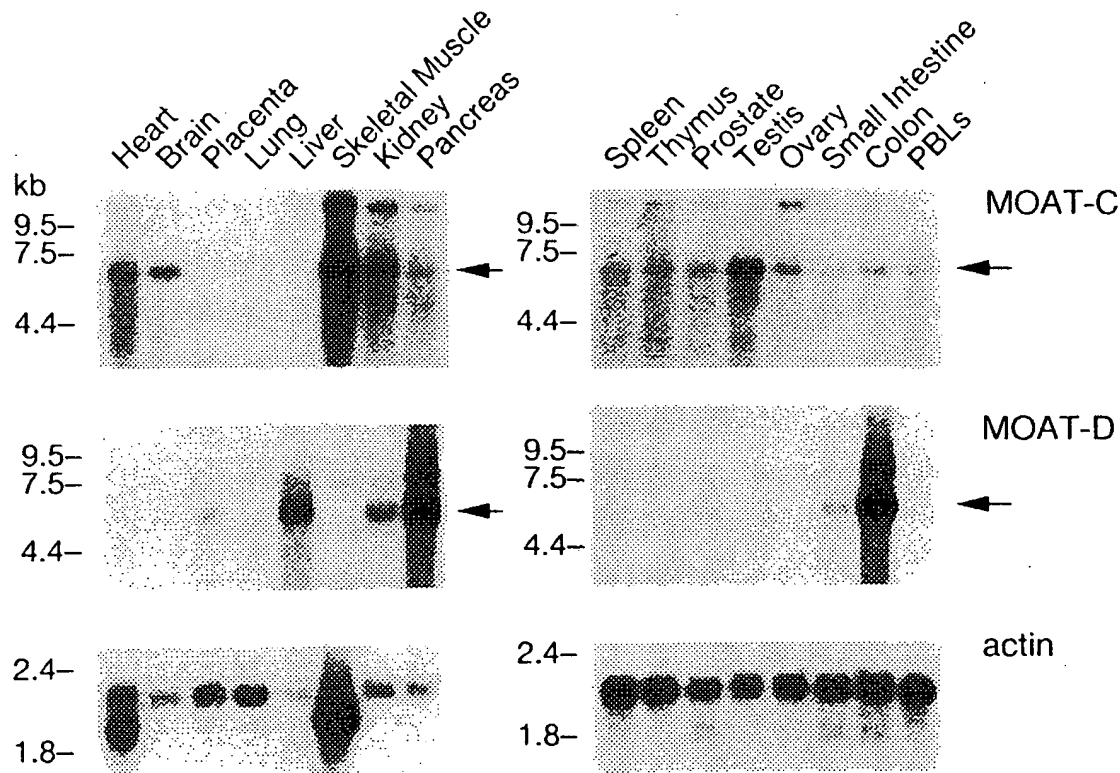


Figure 7

10/56

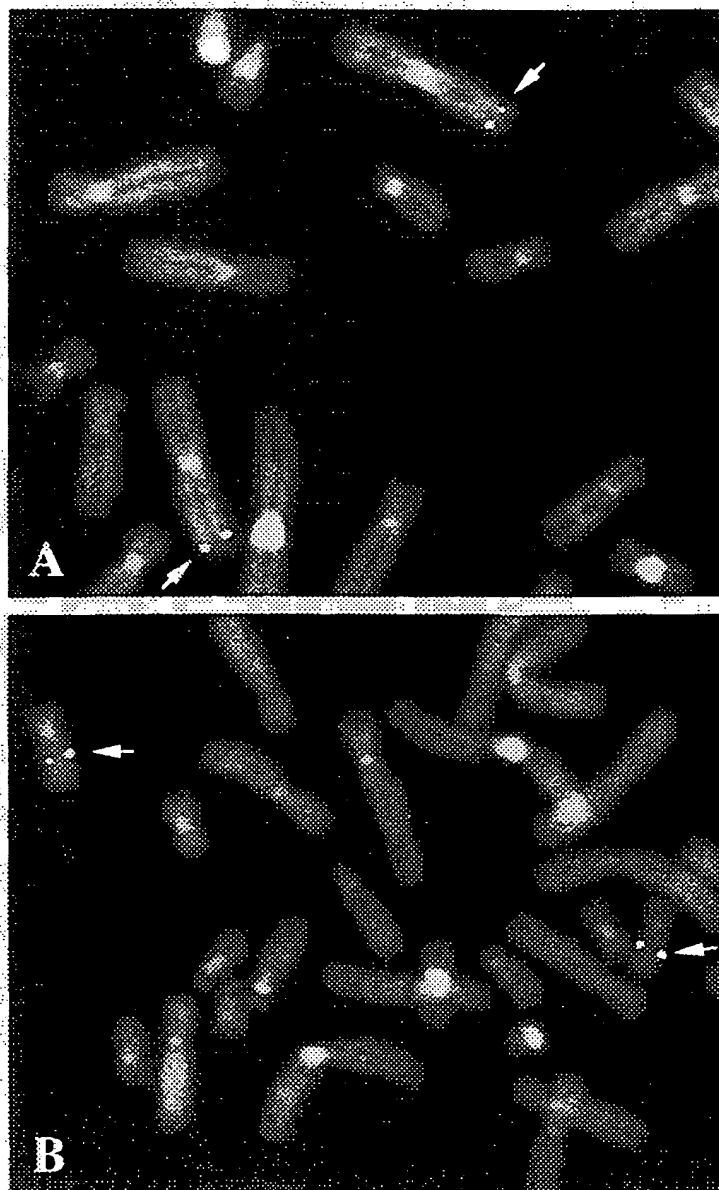


Figure 8

11/56

1 MAAPAEPCAG QGVWNOTEPE PAATSLLSLC FLRTAGVWVP PMYLWVLGPI YLLFIHHHGR  
 61 GYLMSPLFK AKMVLGFALI VLCTSSVAVA LWKIQOQTPE APEFLIHPTV WLTTMSFAVF  
 121 LIHTERKKGV QSSGVLFYW LLCFVLPATN AAQOASGAGF QSDPVRHLST YLCLSLVVAQ  
 181 FVLSCLADQP PFFPEDPQOS NPCPETGAAP PSKATFWWVS GLVWRGYRRP LRPKDLWSLG  
 241 RENSSEELVS RLEKEWMNR SAARRHNKAI AFKRKGGSGM KAPETEPFLR QEGSQWRPLL  
 301 KAIQVVFHST FLLGTLSLII SDVFRFTVPK LLSLFLEFIG DPKPPANKGY LLAVLMFLSA  
 361 CLQTLFEQQN MYRLKVPQMR LRSAITGLVY RKVLALSSGS RKASAVGDVV NLVSVDVQRL  
 421 TESVLYLNL WLPLVWIVC FVYLWQLLGP SALTAIAVFL SLLPLNFFIS KKRNNHQEEQ  
 481 MRQKDSRRL TSSILRNSKT IKFHGWEGAF LDRVLGIRGQ ELGALRTSGL LFSVSLVSFO  
 541 VSTFLVALVV FAVHTLVAEN AMNAEKAFVT LTVLNILNKA QAFLPFSIHS LVQARVSFDR  
 601 LVTFLCLEEV DPGVVDSSSS GSAAGKDCIT IHSATFAWSQ ESPPCLHRIN LTVPOGCCLLA  
 661 VVGPGVAGKS SLLSALLGEL SKVEGFVSIE GAVAYVPQEA WVQNTSVVEN VCFGQELDPP  
 721 WLERVLEACA LQPDVDSFPE GIHTSIGEQG MNLSGGOKOR LSLARAVYRK AAVYLLDDPL  
 781 AALDAHVGQH VFNQVIGPGG LLQGTTRILV THALHILPQA DWIIVLANGA IAEMGSYQEL  
 841 LQRKGALVCL LDQARQPGDR GEGETEPGTS TKDPRGTSAG RRPRLRRERS IKSVPKDR  
 901 TSEAQTEVPL DDPDRAGWPA GKDSIQYGRV KATVHLAYLR AVGTPLCLYA LFLFLCQOVA  
 961 SFRCGYWLSL WADDPVGGQ QTQAALRGGI FGLLGCLQAI GLFASMAAVL LGGARASRL  
 1021 FORLLWDVVR SPISFFERTP IGHLLNRFSK ETDTVDDIP DKLRSLMYA FGLLEVSLVV  
 1081 AVATPLATVA ILPLFLLYAG FQSLYVVSSC QLRRLESASY SSVCSHMAET FQGSTVVR  
 1141 RTOAPFVAQN NARVDESORI SPRLVADRW LAANVELLGN GLVFAAATCA VLSKAHLSAG  
 1201 LVGFSVSAAL QVTQALQWVV RNWTDLENSI VSVERMQDYA WTPKEAPWRL PTCAAQPPWP  
 1261 QGGQIEFRDF GLRYRPELPL AVQGVSLKIH AGEKVGIVGR TGAGKSSLAS GLLRLQEA  
 1321 GGIWIDGVPI AHVGLHTLRS RISIIPQDPI LFPGLRMNL DLLQEHSDA IWAALQTVQL  
 1381 KALVASLPGQ LQYKADRGE DLSVGQKQLL CLARALLRKT QILILDEATA AVDPGTELOM  
 1441 QAMLGSWFAQ CTVLLIAHRL RSVMDCARVL VMDKGQVAES GSPAQLLAQK GLFYRLAQES  
 1501 GLV

Figure 9

12/56

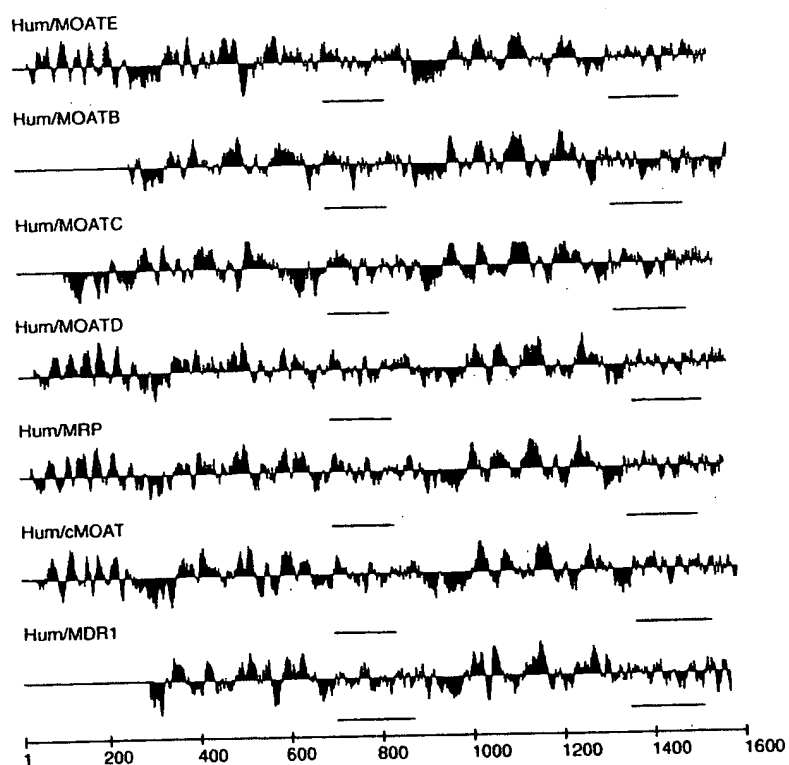


Figure 10

13/56

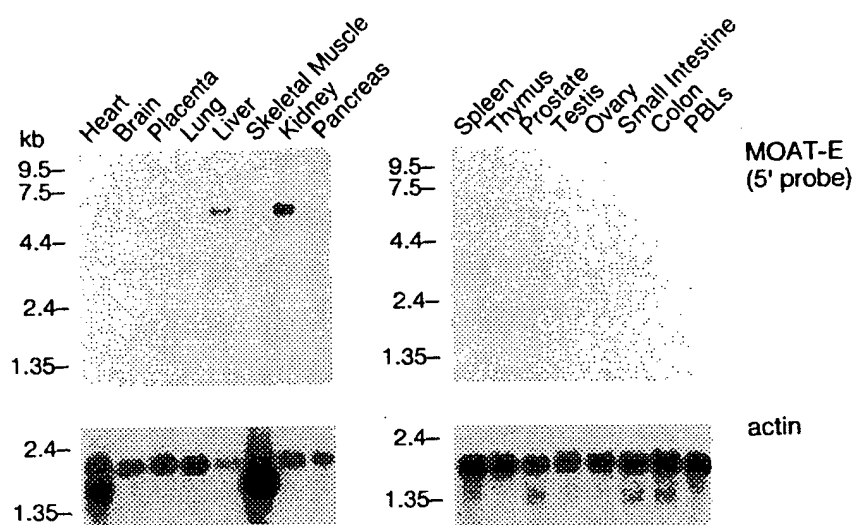


Figure 11

14/56

## MOAT B cDNA AND AMINO ACID SEQUENCE ENCODED THEREBY

ATGCTGCCCGTG TACCAGGAGGTGAAGCCCAACCCGCTGCAGGACGCGAACATCTGCTCA  
1 -----+-----+-----+-----+-----+-----+ 60  
TACGACGGGCACATGGTCCTCCACTTCGGGTTGGGCGACGTCCTGCGCTTGTAGACGAGT

a M L P V Y Q E V K P N P L G D A N I C S -

CGCGTGTTCTTCTGGTGGCTCAATCCCTTGTTTAAAATTGGCCATAAACGGAGATTAGAG  
61 -----+-----+-----+-----+-----+-----+ 120  
GCGCACAAGAAGACCACCGAGTTAGGGAACAAATTTTAACCGGTATTTGCCTCTAATCTC

a R V F F W W L N P L F K I G H K R R L E -

GAAGATGATATGTATTCAGTGCTGCCAGAAGACCGCTCACAGCACCTTGGAGAGGAGTTG  
121 -----+-----+-----+-----+-----+-----+ 180  
CTTCTACTATACATAAGTCACGACGGTCTTCTGGCGAGTGTCGTGGAACCTCTCCTCAAC

a E D D M Y S V L P E D R S Q H L G E E L -

CAAGGGTCTGGGATAAAGAAGTTTTAAGAGCTGAGAATGACGCACAGAAGCCTTCTTTA  
181 -----+-----+-----+-----+-----+-----+ 240  
GTTCCCAAGACCCTATTTCTTCAAAATCTCGACTCTTACTGCGTGTCTTCGGAAGAAAT

a Q G F W D K E V L R A E N D A Q K P S L -

ACAAGAGCAATCATAAAGTGTTACTGGAAATCTTATTTAGTTTTGGGAATTTTACGTTA  
241 -----+-----+-----+-----+-----+-----+ 300  
TGTTCTCGTTAGTATTTACAATGACCTTTAGAATAAATCAAAACCCTTAAAAATGCAAT

a T R A I I K C Y W K S Y L V L G I F T L -

ATTGAGGAAAAGTGCCAAAGTAATCCAGCCCATATTTTTGGGAAAAATTATTAATTATTTT  
301 -----+-----+-----+-----+-----+-----+ 360  
TAACTCCTTTCACGGTTTCATTAGGTCGGGTATAAAAACCCTTTTTAATAATTAATAAAA

a I E E S A K V I Q P I F L G K I I N Y F -

GAAAATTATGATCCCATGGATTCTGTGGCTTTGAACACAGCGTACGCCTATGCCACGGTG

Figure 12A

15/56

361 -----+-----+-----+-----+-----+-----+ 420  
CTTTAATACTAGGGTACCTAAGACACCGAAACTTGTGTGCGCATGCGGATACGGTGCCAC

a E N Y D P M D S V A L N T A Y A Y A T V -

CTGACTTTTTGCACGCTCATTTTGGCTATACTGCATCACTTATATTTTATCACGTTTCAG

421 -----+-----+-----+-----+-----+-----+ 480  
GACTGAAAAACGTGCGAGTAAACCGATATGACGTAGTGAATATAAAAATAGTGCAAGTC

a L T F C T L I L A I L H H L Y F Y H V Q -

TGTGCTGGGATGAGGTTACGAGTAGCCATGTGCCATATGATTTATCGGAAGGCACTTCGT

481 -----+-----+-----+-----+-----+-----+ 540  
ACACGACCCTACTCCAATGCTCATCGGTACACGGTATACTAAATAGCCTTCCGTGAAGCA

a C A G M R L R V A M C H M I Y R K A L R -

CTTAGTAACATGGCCATGGGGAAGACAACCACAGGCCAGATAGTCAATCTGCTGTCCAAT

541 -----+-----+-----+-----+-----+-----+ 600  
GAATCATTGTACCGGTACCCCTTCTGTTGGTGTCCGGTCTATCAGTTAGACGACAGGTTA

a L S N M A M G K T T T G Q I V N L L S N -

GATGTGAACAAGTTTGATCAGGTGACAGTGTCTTACACTTCCTGTGGGCAGGACCACTG

601 -----+-----+-----+-----+-----+-----+ 660  
CTACACTTGTTCAAACCTAGTCCACTGTCACAAGAATGTGAAGGACACCCGTCCTGGTGAC

a D V N K F D Q V T V F L H F L W A G P L -

CAGGCGATCGCAGTGAAGTGCCTACTCTGGATGGAGATAGGAATATCGTGCCTTGCTGGG

661 -----+-----+-----+-----+-----+-----+ 720  
GTCCGCTAGCGTCACTGACGGGATGAGACCTACCTCTATCCTTATAGCACGGAACGACCC

a Q A I A V T A L L W M E I G I S C L A G -

ATGGCAGTTCTAATCATTCTCCTGCCCTTGCAAAGCTGTTTTGGGAAGTTGTTCTCATCA

721 -----+-----+-----+-----+-----+-----+ 780  
TACCGTCAAGATTAGTAAGAGGACGGGAACGTTTCGACAAAACCCTTCAACAAGAGTAGT

a M A V L I I L L P L Q S C F G K L F S S -

CTGAGGAGTAAACCTGCAACTTTCACGGATGCCAGGATCAGGACCATGAATGAAGTTATA

781 -----+-----+-----+-----+-----+-----+ 840



16/56

GACTCCTCATTTTGACGTTGAAAGTGCCTACGGTCCTAGTCCTGGTACTTACTTCAATAT

a L R S K T A T F T D A R I R T M N E V I -

ACTGGTATAAGGATAATAAAAATGTACGCCTGGGAAAAGTCATTTTCAAATCTTATTACC  
841 -----+-----+-----+-----+-----+-----+ 900  
TGACCATATTCCTATTATTTTACATGCGGACCCTTTTCAGTAAAAGTTTAGAATAATGG

a T G I R I I K M Y A W E K S F S N L I T -

AATTTGAGAAAGAAGGAGATTTCCAAGATTCTGAGAAGTTCCTGCCTCAGGGGGATGAAT  
901 -----+-----+-----+-----+-----+-----+ 960  
TTAAACTCTTTCTTCTCTAAAGGTTCTAAGACTCTTCAAGGACGGAGTCCCCCTACTTA

a N L R K K E I S K I L R S S C L R G M N -

TTGGCTTCGTTTTTTCAGTGCAAGCAAAATCATCGTGTGTTGTGACCTTCACCACCTACGTG  
961 -----+-----+-----+-----+-----+-----+ 1020  
AACCGAAGCAAAAAGTCACGTTTCGTTTTAGTAGCACAAACACTGGAAGTGGTGGATGCAC

a L A S F F S A S K I I V F V T F T T Y V -

CTCCTCGGCAGTGTGATCACAGCCAGCCGCGTGTTCGTGGCAGTGACGCTGTATGGGGCT  
1021 -----+-----+-----+-----+-----+-----+ 1080  
GAGGAGCCGTCACACTAGTGTCTGGTCTGGCGCACAAAGCACCGTCACTGCGACATACCCCGA

a L L G S V I T A S R V F V A V T L Y G A -

GTGCGGCTGACGGTTACCTCTTCTTCCCCTCAGCCATTGAGAGGGTGTGAGAGGCAATC  
1081 -----+-----+-----+-----+-----+-----+ 1140  
CACGCCGACTGCCAATGGGAGAAGAAGGGGAGTCGGTAACTCTCCACAGTCTCCGTTAG

a V R L T V T L F F P S A I E R V S E A I -

GTCAGCATCCGAAGAATCCAGACCTTTTTGCTACTTGATGAGATATCACAGCGCAACCGT  
1141 -----+-----+-----+-----+-----+-----+ 1200  
CAGTCGTAGGCTTCTTAGGTCTGGAAAAACGATGAACTACTCTATAGTGTCTGCGTTGGCA

a V S I R R I Q T F L L L D E I S Q R N R -

CAGCTGCCGTCAGATGGTAAAAAGATGGTGCATGTGCAGGATTTTACTGCTTTTGGGAT  
1201 -----+-----+-----+-----+-----+-----+ 1260  
GTCGACGGCAGTCTACCATTTTCTACCACGTACACGTCCTAAAATGACGAAAAACCCTA

17/56

a Q L P S D G K K M V H V Q D F T A F W D -  
AAGGCATCAGAGACCCCAACTCTACAAGGCCTTTCCTTTACTGTCAGACCTGGCGAATTG  
1261 -----+-----+-----+-----+-----+-----+ 1320  
TTCCGTAGTCTCTGGGGTTGAGATGTTCCGGAAGGAAATGACAGTCTGGACCGCTTAAC

a K A S E T P T L O G L S F T V R P G E L -  
TTAGCTGTGGTCGGCCCCGTGGGAGCAGGGAAGTCATCACTGTTAAGTGCCGTGCTCGGG  
1321 -----+-----+-----+-----+-----+-----+ 1380  
AATCGACACCAGCCGGGGCACCCTCGTCCCTTCAGTAGTGACAATTCACGGCAGAGCCC

a L A V V G P V G A G K S S L L S A V L G -  
GAATTGGCCCCAAGTCACGGGCTGGTCAGCGTGCATGGAAGAATTGCCTATGTGTCTCAG  
1381 -----+-----+-----+-----+-----+-----+ 1440  
CTTAACGGGGTTCAGTGCCCGACCAGTCGCACGTACCTTCTTAACGGATACACAGAGTC

a E L A P S H G L V S V H G R I A Y V S Q -  
CAGCCCTGGGTGTTCTCGGGAAGTCTGAGGAGTAATATTTTATTTGGGAAGAAATATGAA  
1441 -----+-----+-----+-----+-----+-----+ 1500  
GTCGGGACCCACAAGAGCCCTTGAGACTCCTCATTATAAAATAAACCCCTTCTTTATACTT

a Q P W V F S G T L R S N I L F G K K Y E -  
AAGGAACGATATGAAAAAGTCATAAAGGCTTGCTCTGAAAAAGGATTTACAGCTGTTG  
1501 -----+-----+-----+-----+-----+-----+ 1560  
TTCCTTGCTATACTTTTTCAGTATTTCCGAACACGAGACTTTTTCCTAAATGTCGACAAC

a K E R Y E K V I K A C A L K K D L Q L L -  
GAGGATGGTGATCTGACTGTGATAGGAGATCGGGGAACACGCTGAGTGAGGGGCAGAAA  
1561 -----+-----+-----+-----+-----+-----+ 1620  
CTCCTACCACTAGACTGACACTATCCTCTAGCCCTTGGTGCGACTCACCTCCCGTCTTT

a E D G D L T V I G D R G T T L S G G Q K -  
GCACGGGTAAACCTTGCAAGAGCAGTGTATCAAGATGCTGACATCTATCTCCTGGACGAT  
1621 -----+-----+-----+-----+-----+-----+ 1680  
CGTGCCCATTTGGAACGTTCTCGTCACATAGTTCTACGACTGTAGATAGAGGACCTGCTA

Figure 12D

18/56

a A R V N L A R A V Y Q D A D I Y L L D D -

CCTCTCAGTGCAGTAGATGCGGAAGTTAGCAGACACTTGTTGGAAGTGTGATTTGTCAA  
1681 -----+-----+-----+-----+-----+-----+ 1740  
GGAGAGTCACGTCATCTACGCCTTCAATCGTCTGTGAACAAGCTTGACACATAAACAGTT

a P L S A V D A E V S R H L F E L C I C O -

ATTTTGCATGAGAAGATCACAATTTTAGTGACTCATCAGTTGCAGTACCTCAAAGCTGCA  
1741 -----+-----+-----+-----+-----+-----+ 1800  
TAAACGTAAGTCTTCTAGTGTTAAATCACTGAGTAGTCAACGTCATGGAGTTTCGACGT

a I L H E K I T I L V T H Q L Q Y L K A A -

AGTCAGATTCTGATATTGAAAGATGGTAAAATGGTGCAGAAGGGGACTTACACTGAGTTC  
1801 -----+-----+-----+-----+-----+-----+ 1860  
TCAGTCTAAGACTATAACTTTCTACCATTTTACCACGTCTTCCCCTGAATGTGACTCAAG

a S Q I L I L K D G K M V Q K G T Y T E F -

CTAAAATCTGGTATAGATTTTGGCTCCCTTTTAAAGAAGGATAATGAGGAAAGTGAACAA  
1861 -----+-----+-----+-----+-----+-----+ 1920  
GATTTTAGACCATATCTAAAACCGAGGGGAAAATTTCTTCTATTACTCCTTTCAGTTGTT

a L K S G I D F G S L L K K D N E E S E Q -

CCTCCAGTTCCAGGAAGTCCCACACTAAGGAATCGTACCTTCTCAGAGTCTTCGGTTTGG  
1921 -----+-----+-----+-----+-----+-----+ 1980  
GGAGGTCAAGGTCCTTGAGGGTGTGATTCTTAGCATGGAAGAGTCTCAGAAGCCAAACC

a P P V P G T P T L R N R T F S E S S V W -

TCTCAACAATCTTCTAGACCCTCCTTGAAAGATGGTGCTCTGGAGAGCCAAGATACAGAG  
1981 -----+-----+-----+-----+-----+-----+ 2040  
AGAGTTGTTAGAAGATCTGGGAGGAAGTCTTCTACCACGAGACCTCTCGGTTCTATGTCTC

a S Q Q S S R P S L K D G A L E S Q D T E -

AATGTCCCAGTTACACTATCAGAGGAGAACCGTTCTGAAGGAAAAGTTGGTTTTTCAGGCC  
2041 -----+-----+-----+-----+-----+-----+ 2100  
TTACAGGGTCAATGTGATAGTCTCCTCTTGGCAAGACTTCTTTTCAACCAAAAGTCCGG

a N V P V T L S E E N R S E G K V G F Q A

19/56

TATAAGAATTACTTCAGAGCTGGTGCTCACTGGATTGTCTTCATTTTCCTTATTCTCCTA  
2101 -----+-----+-----+-----+-----+-----+ 2160  
ATATTCTTAATGAAGTCTCGACCACGAGTGACCTAACAGAAGTAAAAGGAATAAGAGGAT

a Y K N Y F R A G A H W I V F I F L I L L -

AACACTGCAGCTCAGGTTGCCTATGTGCTTCAAGATTGGTGGCTTTCATACTGGGCAAAC  
2161 -----+-----+-----+-----+-----+-----+ 2220  
TTGTGACGTCGAGTCCAACGGATACACGAAGTTCTAACCACCGAAAGTATGACCCGTTTG

a N T A A Q V A Y V L Q D W W L S Y W A N -

AAACAAAGTATGCTAAATGTCACTGTAAATGGAGGAGGAAATGTAACCGAGAAGCTAGAT  
2221 -----+-----+-----+-----+-----+-----+ 2280  
TTTGTTCATACGATTTACAGTGACATTTACCTCCTCCTTTACATTGGCTCTTCGATCTA

a K Q S M L N V T V N G G G N V T E K L D -

CTTAAGTGGTACTTAGGAATTTATTCAGGTTTAACTGTAGCTACCGTTCTTTTGGCATA  
2281 -----+-----+-----+-----+-----+-----+ 2340  
GAATTGACCATGAATCCTTAAATAAGTCCAAATTGACATCGATGGCAAGAAAAACCGTAT

a L N W Y L G I Y S G L T V A T V L F G I -

GCAAGATCTCTATTGGTATTCTACGTCCTTGTTAACTCTTCACAAACTTTGCACAACAAA  
2341 -----+-----+-----+-----+-----+-----+ 2400  
CGTTCTAGAGATAACCATAAGATGCAGGAACAATTGAGAAGTGTTTGAAACGTGTTGTTT

a A R S L L V F Y V L V N S S Q T L H N K -

ATGTTTGAGTCAATTCTGAAAGCTCCGGTATTATTCTTTGATAGAAATCCAATAGGAAGA  
2401 -----+-----+-----+-----+-----+-----+ 2460  
TACAAACTCAGTTAAGACTTTCGAGGCCATAATAAGAACTATCTTTAGGTTATCCTTCT

a M F E S I L K A P V L F F D R N P I G R -

ATTTTAAATCGTTTCTCCAAAGACATTGGACACTTGGATGATTTGCTGCCGCTGACGTTT  
2461 -----+-----+-----+-----+-----+-----+ 2520  
TAAATTTAGCAAAGAGGTTTCTGTAACCTGTGAACCTACTAAACGACGGCGACTGCAAA

a I L N R F S K D I G H L D D L L P L T F

Figure 12F

20/56

TTAGATTTTCATCCAGACATTGCTACAAGTGGTTGGTGTGGTCTCTGTGGCTGTGGCCGTG  
2521 -----+-----+-----+-----+-----+-----+ 2580  
AATCTAAAGTAGGTCTGTAACGATGTTACCAACCACACCAGAGACACCGACACCGGCAC

a L D F I Q T L L Q V V G V V S V A V A V -

ATTCTTGGATCGCAATACCCTTGGTTCCCCTTGAATCATTTTCATTTTCTTCGGCGA  
2581 -----+-----+-----+-----+-----+-----+ 2640  
TAAGGAACCTAGCGTTATGGAACCAAGGGGAACCTTAGTAAAAGTAAAAAGAAGCCGT

a I P W I A I P I V P L G I I F I F L R R -

TATTTTTTGGAAACGTCAAGAGATGTGAAGCGCCTGGAATCTACAACTCGGAGTCCAGTG  
2641 -----+-----+-----+-----+-----+-----+ 2700  
ATAAAAAACCTTGCAGTTCTCTACACTTCGCGGACCTTAGATGTTGAGCCTCAGGTCAC

a Y F L E T S R D V K R L E S T T R S P V -

TTTTCCCACTTGTCATCTTCTCTCCAGGGGCTCTGGACCATCCGGGCATACAAAGCAGAA  
2701 -----+-----+-----+-----+-----+-----+ 2760  
AAAAGGGTGAACAGTAGAAGAGAGGTCCCCGAGACCTGGTAGGCCCGTATGTTTCGTCTT

a F S H L S S S L Q G L W T I R A Y K A E -

GAGAGGTGTCAGGAACTGTTTGATGCACACCAGGATTTACATTCAGAGGCTTGGTTCTTG  
2761 -----+-----+-----+-----+-----+-----+ 2820  
CTCTCCACAGTCCTTGACAACTACGTGTGGTCCTAAATGTAAGTCTCCGAACCAAGAAC

a E R C Q E L F D A H Q D L H S E A W F L -

TTTTTGACAACGTCCCGCTGGTTCGCCGTCCGTCTGGATGCCATCTGTGCCATGTTTGTC  
2821 -----+-----+-----+-----+-----+-----+ 2880  
AAAAACTGTTGCAGGGCGACCAAGCGGCAGGCAGACCTACGGTAGACACGGTACAAACAG

a F L T T S R W F A V R L D A I C A M F V -

ATCATCGTTGCCTTTGGGTCCCTGATTCTGGCAAAAACCTCTGGATGCCGGGCAGGTTGGT  
2881 -----+-----+-----+-----+-----+-----+ 2940  
TAGTAGCAACGGAAACCCAGGGACTAAGACCGTTTTTGAGACCTACGGCCCGTCCAACCA

a I I V A F G S L I L A K T L D A G Q V G -

TTGGCACTGTCCTATGCCCTCACGCTCATGGGGATGTTTCAGTGGTGTGTTGACAAAGT

Figure 12G

21/56

2941 -----+-----+-----+-----+-----+-----+ 3000  
AACCGTGACAGGATACGGGAGTGCGAGTACCCCTACAAAGTCACCACACAAGCTGTTTCA

a L A L S Y A L T L M G M F Q W C V R Q S -

GCTGAAGTTGAGAATATGATGATCTCAGTAGAAAGGGTCATTGAATACACAGACCTTGAA  
3001 -----+-----+-----+-----+-----+-----+ 3060  
CGACTTCAACTCTTATACTACTAGAGTCATCTTTCCCAGTAACTTATGTGTCTGGAACCT

a A E V E N M M I S V E R V I E Y T D L E -

AAAGAAGCACCTTGGAATATCAGAAACGCCCACCACCAGCCTGGCCCCATGAAGGAGTG  
3061 -----+-----+-----+-----+-----+-----+ 3120  
TTTCTTCGTGGAACCCTTATAGTCTTTGCGGGTGGTGGTCGGACCGGGGTACTTCCTCAC

a K E A P W E Y Q K R P P P A W P H E G V -

ATAATCTTTGACAATGTGAACTTCATGTACAGTCCAGGTGGGCCTCTGGTACTGAAGCAT  
3121 -----+-----+-----+-----+-----+-----+ 3180  
TATTAGAAACTGTTACACTTGAAGTACATGTCAGGTCCACCCGGAGACCATGACTTCGTA

a I I F D N V N F M Y S P G G P L V L K H -

CTGACAGCACTCATTAAATCACAAGAAAAGGTTGGCATTGTGGGAAGAACCGGAGCTGGA  
3181 -----+-----+-----+-----+-----+-----+ 3240  
GACTGTCGTGAGTAATTTAGTGTTCTTTTCCAACCGTAACACCCTTCTTGGCCTCGACCT

a L T A L I K S Q E K V G I V G R T G A G -

AAAAGTTCCTCATCTCAGCCCTTTTATAGATTGTCAGAACCCGAAGGTAAAATTTGGATT  
3241 -----+-----+-----+-----+-----+-----+ 3300  
TTTTCAAGGGAGTAGAGTCGGGAAAAATCTAACAGTCTTGGGCTTCCATTTTAAACCTAA

a K S S L I S A L F R L S E P E G K I W I -

GATAAGATCTTGACAACTGAAATTGGACTTCACGATTTAAGGAAGAAAATGTCAATCATA  
3301 -----+-----+-----+-----+-----+-----+ 3360  
CTATTCTAGAACTGTTGACTTTAACCTGAAGTGCTAAATTCCTTCTTTTACAGTTAGTAT

a D K I L T T E I G L H D L R K K M S I I -

CCTCAGGAACCTGTTTTGTTCACTGGAACAATGAGGAAAAACCTGGATCCCTTTAAGGAG  
3361 -----+-----+-----+-----+-----+-----+ 3420

Figure 12H

22/56

GGAGTCCTTGGACAAAACAAGTGACCTTGTTACTCCTTTTTGGACCTAGGGAAATTCCTC

a P Q E P V L F T G T M R K N L D P F K E -

CACACGGATGAGGAACTGTGGAATGCCTTACAAGAGGTACAACCTTAAAGAAACCATTGAA

3421 -----+-----+-----+-----+-----+-----+ 3480

GTGTGCCTACTCCTTGACACCTTACGGAATGTTCTCCATGTTGAATTTCTTTGGTAACTT

a H T D E E L W N A L Q E V Q L K E T I E -

GATCTTCCTGGTAAAATGGATACTGAATTAGCAGAATCAGGATCCAATTTTAGTGTTGGA

3481 -----+-----+-----+-----+-----+-----+ 3540

CTAGAAGGACCATTTTACCTATGACTTAATCGTCTTAGTCCTAGGTTAAAATCACAACT

a D L P G K M D T E L A E S G S N F S V G -

CAAAGACAACCTGGTGTGCCTTGCCAGGGCAATTCTCAGGAAAAATCAGATATTGATTATT

3541 -----+-----+-----+-----+-----+-----+ 3600

GTTTCTGTTGACCACACGGAACGGTCCCCTTAAGAGTCCTTTTTTAGTCTATACTAATAA

a Q R Q L V C L A R A I L R K N Q I L I I -

GATGAAGCGACGGCAAATGTGGATCCAAGAACTGATGAGTTAATACAAAAAAAAATCCGG

3601 -----+-----+-----+-----+-----+-----+ 3660

CTACTTCGCTGCCGTTTACACCTAGGTTCTTGACTACTCAATTATGTTTTTTTTTAGGCC

a D E A T A N V D P R T D E L I Q K K I R -

GAGAAATTTGCCCACTGCACCGTGCTAACCATTGCACACAGATTGAACACCATTATTGAC

3661 -----+-----+-----+-----+-----+-----+ 3720

CTCTTTAAACGGGTGACGTGGCAGGATTGGTAACGTGTGTCTAACTTGTGGTAATAACTG

a E K F A H C T V L T I A H R L N T I I D -

AGCGACAAGATAATGGTTTTAGATTACAGGAAGACTGAAAGAATATGATGAGCCGTATGTT

3721 -----+-----+-----+-----+-----+-----+ 3780

TCGCTGTTCTATTACCAAATCTAAGTCCTTCTGACTTTCTTATACTACTCGGCATACAA

a S D K I M V L D S G R L K E Y D E P Y V -

TTGCTGCAAAATAAAGAGAGCCTATTTTACAAGATGGTGCAACAACCTGGGCAAGGCAGAA

3781 -----+-----+-----+-----+-----+-----+ 3840

AACGACGTTTTATTTCTCTCGGATAAAATGTTCTACCACGTTGTTGACCCGTTCCGTCTT

23/56

a L L Q N K E S L F Y K M V Q Q L G K A E .

GCCGCTGCCCTCACTGAAACAGCAAAACAGGTATACTTCAAAAGAAATTATCCACATATT  
3841 -----+-----+-----+-----+-----+-----+ 3900  
CGGCGACGGGAGTGACTTTGTCGTTTTGTCCATATGAAGTTTTCTTTAATAGGTGTATAA

a A A A L T E T A K Q V Y F K R N Y P H I .

GGTCACACTGACCACATGGTTACAAACACTTCCAATGGACAGCCCTCGACCTTAACTATT  
3901 -----+-----+-----+-----+-----+-----+ 3960  
CCAGTGTGACTGGTGTACCAATGTTTGTGAAGGTTACCTGTCGGGAGCTGGAATTGATAA

a G H T D H M V T N T S N G Q P S T L T I .

TTCGAGACAGCACTG  
3961 -----+----- 3975  
AAGCTCTGTCGTGAC

a F E T A L .

Figure 12J



24/56

## MOAT C cDNA AND AMINO ACID SEQUENCE ENCODED THEREBY

ATGAAGGATATCGACATAGGAAAAGAGTATATCATCCCCAGTCCTGGGTATAGAAGTGTG  
 1 -----+-----+-----+-----+-----+-----+ 60  
 TACTTCCTATAGCTGTATCCTTTTCTCATATAGTAGGGGTCAGGACCCATATCTTCACAC  
 a M K D I D I G K E Y I I P S P G Y R S V -

AGGGAGAGAACCAGCACTTCTGGGACGCACAGAGACCGTGAAGATTCCAAGTTCAGGAGA  
 61 -----+-----+-----+-----+-----+-----+ 120  
 TCCCTCTCTTGGTCGTGAAGACCCTGCGTGTCTCTGGCACTTCTAAGGTTCAAGTCCTCT  
 a R E R T S T S G T H R D R E D S K F R R -

ACTCGACCGTTGGAATGCCAAGATGCCTTGAAACAGCAGCCCGAGCCGAGGGCCTCTCT  
 121 -----+-----+-----+-----+-----+-----+ 180  
 TGAGCTGGCAACCTTACGGTTCTACGGAACCTTTGTCGTCGGGCTCGGCTCCCGGAGAGA  
 a T R P L E C Q D A L E T A A R A E G L S -

CTTGATGCCTCCATGCATTCTCAGCTCAGAATCCTGGATGAGGAGCATCCCAAGGGAAAG  
 181 -----+-----+-----+-----+-----+-----+ 240  
 GAACTACGGAGGTACGTAAGAGTCGAGTCTTAGGACCTACTCCTCGTAGGGTTCCCTTTC  
 a L D A S M H S Q L R I L D E E H P K G K -

TACCATCATGGCTTGAGTGCTCTGAAGCCCATCCGGACTACTTCCAAACACCAGCACCCA  
 241 -----+-----+-----+-----+-----+-----+ 300  
 ATGGTAGTACCGAACTCACGAGACTTCGGGTAGGCCTGATGAAGGTTTGTGGTCGTGGGT  
 a Y H H G L S A L K P I R T T S K H Q H P -

GTGGACAATGCTGGGCTTTTTTCTGTATGACTTTTTCTGGGCTTTCTTCTCTGGCCCGT  
 301 -----+-----+-----+-----+-----+-----+ 360  
 CACCTGTTACGACCCGAAAAAAGGACATACTGAAAAAGCACCGAAAGAAGAGACCGGGCA  
 a V D N A G L F S C M T F S W L S S L A R -

GTGGCCCAACAAGAAGGGGGAGCTCTCAATGGAAGACGTGTGGTCTCTGTCCAAGCACGAG

Figure 13A

25/56

361 -----+-----+-----+-----+-----+-----+ 420  
CACCGGGTGTCTTCCCCCTCGAGAGTTACCTTCTGCACACCAGAGACAGGTTCTGTGCTC

a V A H K K G E L S M E D V W S L S K H E -

TCTTCTGACGTGAACTGCAGAAGACTAGAGAGACTGTGGCAAGAAGAGCTGAATGAAGTT

421 -----+-----+-----+-----+-----+-----+ 480  
AGAAGACTGCACTTGACGTCTTCTGATCTCTCTGACACCGTTCTTCTCGACTTACTTCAA

a S S D V N C R R L E R L W Q E E L N E V -

GGGCCAGACGCTGCTTCCCTGCGAAGGGTTGTGTGGATCTTCTGCCGCACCAGGCTCATC

481 -----+-----+-----+-----+-----+-----+ 540  
CCCGGTCTGCGACGAAGGGACGCTTCCCAACACACCTAGAAGACGGCGTGGTCCGAGTAG

a G P D A A S L R R V V W I F C R T R L I -

CTGTCCATCGTGTGCCTGATGATCACGCAGCTGGCTGGCTTCAGTGGACCAGCCTTCATG

541 -----+-----+-----+-----+-----+-----+ 600  
GACAGGTAGCACACGGACTACTAGTGCCTGACCGACCGAAGTCACCTGGTCGGAAGTAC

a L S I V C L M I T Q L A G F S G P A F M -

GTGAAACACCTCTTGGAGTATACCCAGGCAACAGAGTCTAACCTGCAGTACAGCTTGTTG

601 -----+-----+-----+-----+-----+-----+ 660  
CACTTTGTGGAGAACCTCATATGGGTCCGTTGTCTCAGATTGGACGTCATGTGCAACAAC

a V K H L L E Y T Q A T E S N L Q Y S L L -

TTAGTGCTGGGCCTCCTCCTGACGGAAATCGTGCGGTCTTGGTCGCTTGCACTGACTTGG

661 -----+-----+-----+-----+-----+-----+ 720  
AATCACGACCCGGAGGAGGACTGCCTTTAGCACGCCAGAACCAGCGAACGTGACTGAACC

a L V L G L L L T E I V R S W S L A L T W -

GCATTGAATTACCGAACCGGTGTCCGCTTGCGGGGGGCCATCCTAACCATGGCATTTAAG

721 -----+-----+-----+-----+-----+-----+ 780  
CGTAACTTAATGGCTTGGCCACAGGCGAACGCCCCCGGTAGGATTGGTACCGTAAATTC

a A L N Y R T G V R L R G A I L T M A F K -

AAGATCCTTAAGTTAAAGAACAATTAAAGAGAAATCCCTGGGTGAGCTCATCAACATTTGC

781 -----+-----+-----+-----+-----+-----+ 840

Figure 13B

26/56

TTCTAGGAATTCAATTTCTTGTAATTTCTCTTTAGGGACCCACTCGAGTAGTTGTAAACG

a K I L K L K N I K E K S L G E L I N I C .

TCCAACGATGGGCAGAGAATGTTTGAGGCAGCAGCCGTTGGCAGCCTGCTGGCTGGAGGA  
841 -----+-----+-----+-----+-----+-----+ 900

AGGTTGCTACCCGCTCTTACAAACTCCGTCGTCGGCAACCGTCGGACGACCGACCTCCT

a S N D G Q R M F E A A A V G S L L A G G .

CCCGTTGTTGCCATCTTAGGCATGATTTATAATGTAATTATTCTGGGACCAACAGGCTTC  
901 -----+-----+-----+-----+-----+-----+ 960

GGGCAACAACGGTAGAATCCGTACTAAATATTACATTAATAAGACCCTGGTTGTCCGAAG

a P V V A I L G M I Y N V I I L G P T G F .

CTGGGATCAGCTGTTTTATCCTCTTTTACCCAGCAATGATGTTTGCATCACGGCTCACA  
961 -----+-----+-----+-----+-----+-----+ 1020

GACCCTAGTCGACAAAAATAGGAGAAAATGGGTCGTTACTACAAACGTAGTGCCGAGTGT

a L G S A V F I L F Y P A M M F A S R L T .

GCATATTTCAAGGAGAAAATGCGTGGCCGCCACGGATGAACGTGTCCAGAAGATGAATGAA  
1021 -----+-----+-----+-----+-----+-----+ 1080

CGTATAAAGTCCTCTTTTACGCACCGGCGGTGCCTACTTGACAGGTCTTCTACTTACTT

a A Y F R R K C V A A T D E R V Q K M N E .

GTTCTTACTTACATTAAATTTATCAAAATGTATGCCTGGGTCAAAGCATTTTCTCAGAGT  
1081 -----+-----+-----+-----+-----+-----+ 1140

CAAGAATGAATGTAATTTAAATAGTTTTACATACGGACCCAGTTTCGTAAAAGAGTCTCA

a V L T Y I K F I K M Y A W V K A F S Q S .

GTTCAAGAAAATCCGCGAGGAGGAGCGTCGGATATTGGAAAAAGCCGGTACTTCCAGGGT  
1141 -----+-----+-----+-----+-----+-----+ 1200

CAAGTCTTTTAGGCGCTCCTCCTCGCAGCCTATAACCTTTTTCGGCCCATGAAGGTCCCA

a V Q K I R E E E R R I L E K A G Y F Q G .

ATCACTGTGGGTGTGGCTCCCATTGTGGTGGTGATTGCCAGCGTGGTGACCTTCTCTGTT  
1201 -----+-----+-----+-----+-----+-----+ 1260

TAGTGACACCCACACCGAGGGTAACACCACCACTAACGGTCGCACCACTGGAAGAGACAA

27/56

a I T V G V A P I V V V I A S V V T F S V .  
CATATGACCCTGGGCTTCGATCTGACAGCAGCACAGGCTTTCACAGTGGTGACAGTCTTC  
1261 -----+-----+-----+-----+-----+-----+ 1320  
GTATACTGGGACCCGAAGCTAGACTGTCGTCGTGTCCGAAAGTGTCACCACTGTCAGAAG

a H M T L G F D L T A A Q A F T V V T V F .  
AATCCATGACTTTTGCTTTGAAAGTAACACCGTTTTTCAGTAAAGTCCCTCTCAGAAGCC  
1321 -----+-----+-----+-----+-----+-----+ 1380  
TTAAGGTACTGAAAACGAAACTTTCATTGTGGCAAAAGTCATTTTCAGGGAGAGTCTTCGG

a N S M T F A L K V T P F S V K S L S E A .  
TCAGTGGCTGTTGACAGATTTAAGAGTTTGTTTCTAATGGAAGAGGTTACATGATAAAG  
1381 -----+-----+-----+-----+-----+-----+ 1440  
AGTCACCGACAACGTCTAAATTCTCAAACAAAGATTACCTTCTCCAAGTGACTATTTTC

a S V A V D R F K S L F L M E E V H M I K .  
AACAAACCAGCCAGTCCTCACATCAAGATAGAGATGAAAAATGCCACCTTGGCATGGGAC  
1441 -----+-----+-----+-----+-----+-----+ 1500  
TTGTTTGGTCGGTCAGGAGTGTAGTTCTATCTCTACTTTTTACGGTGGAACCGTACCCTG

a N K P A S P H I K I E M K N A T L A W D .  
TCCTCCCACTCCAGTATCCAGAACTCGCCCAAGCTGACCCCCAAAATGAAAAAAGACAAG  
1501 -----+-----+-----+-----+-----+-----+ 1560  
AGGAGGGTGAGGTCATAGGTCTTGAGCGGGTTCGACTGGGGGTTTTACTTTTTCTGTTC

a S S H S S I Q N S P K L T P K M K K D K .  
AGGGCTTCCAGGGGCAAGAAAGAGAAGGTGAGGCAGCTGCAGCGCACTGAGCATCAGGCG  
1561 -----+-----+-----+-----+-----+-----+ 1620  
TCCCGAAGGTCCCCGTTCTTTCTCTTCCACTCCGTCGACGTCGCGTGACTCGTAGTCCGC

a R A S R G K K E K V R Q L Q R T E H Q A .  
GTGCTGGCAGAGCAGAAAGGCCACCTCCTCCTGGACAGTGACGAGCGGCCAGTCCCGAA  
1621 -----+-----+-----+-----+-----+-----+ 1680  
CACGACCGTCTCGTCTTTCCGGTGGAGGAGGACCTGTCACTGCTCGCCGGGTGAGGGCTT

Figure 13D

28/56

a V L A E O K G H L L L D S D E R P S P E .

GAGGAAGAAGGCAAGCACATCCACCTGGGCCACCTGCGCTTACAGAGGACACTGCACAGC  
1681 -----+-----+-----+-----+-----+-----+ 1740  
CTCCTTCTTCCGTTTCGTGTAGGTGGACCCGGTGGACGCGAATGTCTCCTGTGACGTGTCTG

a E E E G K H I H L G H L R L O R T L H S .

ATCGATCTGGAGATCCAAGAGGGTAAACTGGTTGGAATCTGCGGCAGTGTGGGAAGTGGA  
1741 -----+-----+-----+-----+-----+-----+ 1800  
TAGCTAGACCTCTAGGTTCTCCCATTTGACCAACCTTAGACGCCGTCACACCCTTCACCT

a I D L E I Q E G K L V G I C G S V G S G .

AAAACCTCTCTCATTTTCAGCCATTTTAGGCCAGATGACGCTTCTAGAGGGCAGCATTGCA  
1801 -----+-----+-----+-----+-----+-----+ 1860  
TTTTGGAGAGAGTAAAGTCGGTAAAATCCGGTCTACTGCGAAGATCTCCCGTCGTAACGT

a K T S L I S A I L G Q M T L L E G S I A .

ATCAGTGGAACCTTCGCTTATGTGGCCCAGCAGGCCTGGATCCTCAATGCTACTCTGAGA  
1861 -----+-----+-----+-----+-----+-----+ 1920  
TAGTCACCTTGGAAGCGAATACACCGGGTCGTCCGGACCTAGGAGTTACGATGAGACTCT

a I S G T F A Y V A Q Q A W I L N A T L R .

GACAACATCCTGTTTGGAAGGAATATGATGAAGAAAGATACTCTGTGCTGAACAGC  
1921 -----+-----+-----+-----+-----+-----+ 1980  
CTGTTGTAGGACAAACCCTTCCTTATACTACTTCTTTCTATGTTGAGACACGACTTGTCG

a D N I L F G K E Y D E E R Y N S V L N S .

TGCTGCCTGAGGCCTGACCTGGCCATTCTTCCAGCAGCGACCTGACGGAGATTGGAGAG  
1981 -----+-----+-----+-----+-----+-----+ 2040  
ACGACGGACTCCGGACTGGACCGGTAAGAAGGGTCGTCGCTGGACTGCCTCTAACCTCTC

a C C L R P D L A I L P S S D L T E I G E .

CGAGGAGCCAACCTGAGCGGTGGGCAGCGCCAGAGGATCAGCCTTGCCCGGGCCTTGAT  
2041 -----+-----+-----+-----+-----+-----+ 2100  
GCTCCTCGGTTGGACTCGCCACCGTCGCGGTCTCCTAGTCGGAACGGGCCCGGAACATA

a R G A N L S G G O R O R I S L A R A L Y .

29/56

AGTGACAGGAGCATCTACATCCTGGAACGACCCCTCAGTGCCTTAGATGCCCATGTGGGC  
2101 -----+-----+-----+-----+-----+-----+ 2160  
TCACTGTCCTCGTAGATGTAGGACCTGCTGGGGGAGTCACGGAATCTACGGGTACACCCG

a S D R S I Y I L D D P L S A L D A H V G -

AACCACATCTTCAATAGTGCTATCCGGAACATCTCAAGTCCAAGACAGTTCTGTTTGT  
2161 -----+-----+-----+-----+-----+-----+ 2220  
TTGGTGTAAGAAGTTATCACGATAGGCCTTTGTAGAGTTCAGGTTCTGTCAAGACAAACAA

a N H I F N S A I R K H L K S K T V L F V -

ACCCACCAGTTACAGTACCTGGTTGACTGTGATGAAGTGATCTTCATGAAAGAGGGCTGT  
2221 -----+-----+-----+-----+-----+-----+ 2280  
TGGGTGGTCAATGTCATGGACCAACTGACACTACTTCACTAGAAGTACTTTCTCCCGACA

a T H Q L Q Y L V D C D E V I F M K E G C -

ATTACGGAAAGAGGCACCCATGAGGAACTGATGAATTTAAATGGTGACTATGCTACCATT  
2281 -----+-----+-----+-----+-----+-----+ 2340  
TAATGCCTTTCTCCGTGGGTACTCCTTGACTACTTAAATTTACCACTGATACGATGGTAA

a I T E R G T H E E L M N L N G D Y A T I -

TTTAATAACCTGTTGCTGGGAGAGACACCGCCAGTTGAGATCAATTCAAAAAAGGAAACC  
2341 -----+-----+-----+-----+-----+-----+ 2400  
AAATTATTGGACAACGACCCTCTCTGTGGCGGTCAACTCTAGTTAAGTTTTTCTTTGG

a F N N L L L G E T P P V E I N S K K E T -

AGTGGTTCACAGAAGAAGTCACAAGACAAGGGTCCTAAACAGGATCAGTAAAGAAGGAA  
2401 -----+-----+-----+-----+-----+-----+ 2460  
TCACCAAGTGCTTCTTCAGTGTTCTGTTCCAGGATTTTGTCTAGTCATTCTCTCTT

a S G S Q K K S Q D K G P K T G S V K K E -

AAAGCAGTAAAGCCAGAGGAAGGGCAGCTTGTGCAGCTGGAAGAGAAAGGGCAGGGTTCA  
2461 -----+-----+-----+-----+-----+-----+ 2520  
TTTCGTCATTTCCGGTCTCCTTCCCGTCGAACACGTCGACCTTCTCTTTCCCGTCCCAAGT

a K A V K P E E G O L V O L E E K G O G S -

Figure 13F

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**